THESIS / THÈSE

DOCTOR OF SCIENCES

# Perturbed auxiliary problem methods to solve generalized variational inequalities

Salmon, Genevieve

*Award date:*
2001

*Awarding institution:*
University of Namur

**FACULTES UNIVERSITAIRES**

**NOTRE-DAME DE LA PAIX** **NAMUR**

FACULTE DES SCIENCES

# Perturbed Auxiliary Problem Methods to Solve Generalized Variational Inequalities

Dissertation présentée par
**Geneviève SALMON**
pour l'obtention du grade
de Docteur en Sciences

Composition du Jury:

Jean-Jacques Strodiot (Promoteur)
Jean-Pierre Crouzeix
Jacques Henrard
Van Hien Nguyen
Michel Willem

Avril 2001

J'exprime ma reconnaissance envers mon promoteur, le Professeur Jean-Jacques Strodiot. L'aide continue qu'il m'a apportée durant ces années a joué un rôle essentiel dans ce travail. Je le remercie aussi pour sa disponibilité, sa clairvoyance et son sens de l'enseignement.

Ma reconnaissance va aussi au Professeur Van Hien Nguyen pour ses questions constructives et ses conseils avisés.

Je remercie le Professeur Jacques Henrard de présider le jury et les Professeurs Jean-Pierre Crouzeix et Michel Willem de me faire l'honneur de juger ce travail.

Merci aussi à tous ceux qui m'ont aidée dans ma recherche et plus particulièrement, Naïma El-Farouq, Daniel Goeleven, Abdel Hamdi, Sylvianne Haubruge, Panagiotis D. Panagiotopoulos, Georgios E. Stavroulakis, Rainer Tichatschke,...

Mes remerciements vont également aux membres du Département pour leur présence et leurs encouragements.

Je remercie enfin ma famille et mes amis pour tant de choses.

# Contents

Introduction . . . . . . . . . . 3

**1 Background Notes** **11**

1.1 Elements of Convex Analysis . . . . . . . . . . 12

1.2 Preliminaries on Nonlinear Mappings . . . . . . . . . . 14

1.2.1 Continuity and Related Concepts . . . . . . . . . . 15

1.2.2 Derivative of a Singlevalued Mapping . . . . . . . . . . 17

1.2.3 Monotonicity and Related Topics . . . . . . . . . . 18

1.2.4 Dunn Property . . . . . . . . . . 22

1.2.5 Paramonotonicity . . . . . . . . . . 24

1.2.6 The Case of Affine Operators in $I\!R^n$ . . . . . . . . . . 27

1.3 Existence Theory . . . . . . . . . . 28

**2 The Auxiliary Problem Method** **31**

2.1 The Auxiliary Problem Framework and Particular Instances . 31

2.2 Convergence of the Auxiliary Problem Method . . . . . . . . . . 36

2.2.1 The Singlevalued Case . . . . . . . . . . 37

2.2.2 The Multivalued Case . . . . . . . . . . 41

2.3 Projection Methods for Singular and Linear Variational Inequalities in Finite Dimension . . . . . . . . . . 43

**3 Perturbation Methods** **46**

3.1 Convergence of Mosco and Approximation of Variational Inequalities . . . . . . . . . . 46

3.2 The Perturbed Auxiliary Problem . . . . . . . . . . 58

3.3 $\alpha$–Order Convergence . . . . . . . . . . 61

**4 Convergence of the Perturbed Auxiliary Problem Method for Singlevalued Mappings** **64**
4.1 Global Convergence Results . . . . . . . . . . . . . . . . . . . . 65
4.2 Convergence Results for Particular Choices of the Auxiliary Operators . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
4.2.1 The Symmetric Case . . . . . . . . . . . . . . . . . . . . 81
4.2.2 The Nonsymmetric Case . . . . . . . . . . . . . . . . . . 84
4.3 Local Convergence Results. Application to the Newton Method 90

**5 Convergence of the Perturbed Auxiliary Problem Method for Multivalued Mappings** **100**
5.1 Convergence Results . . . . . . . . . . . . . . . . . . . . . . . . 101
5.2 Relaxation of the Scheme and Enlargements of Multivalued Mappings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 118
5.3 Convergence Results for the Inexact Perturbed Auxiliary Problem Method . . . . . . . . . . . . . . . . . . . . . . . . . . . . 124

**6 A Bundle Method to Solve Multivalued Variational Inequalities** **131**
6.1 Approximate Auxiliary Subproblem and Bundle Strategy . . 132
6.2 Convergence of the Algorithm . . . . . . . . . . . . . . . . . . 138

# Introduction

Let $F$ be a monotone multivalued operator defined on a real Hilbert space $H$, and let $\varphi : H \to I\!R \cup \{+\infty\}$ be a lower semi-continuous proper convex function whose effective domain is included in the domain of $F$. The problem considered is the following:

$$(GVIP) \begin{cases} \text{find } x^* \in H \text{ and } r(x^*) \in F(x^*) \text{ such that, for all } x \in H, \\ \langle r(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0. \end{cases}$$

This problem is known as a *generalized variational inequality* (see [48]). It appears in many fields of applied mathematics such as convex programming, partial differential equations, game theory, equilibrium models in economics and transportation sciences, mechanics, physics, optimal control,... We refer to [10], [44], [56], [68], [71], [86], [97], [102], [132] as references sources for the numerous applications of problem $(GVIP)$. Note that problems like $(GVIP)$ with a multivalued mapping $F$ and a function $\varphi$ which is not necessarily the indicator function of a closed convex subset of $H$, are encountered in many applications. In particular, it is the case in mechanical problems as for example, in [98], and in equilibrium problems as, for example, in [32], [92].

By using the definition of the subdifferential of the function $\varphi$, problem $(GVIP)$ can be equivalently written under the following inclusion form:

$$\text{find } x^* \in H : 0 \in F(x^*) + \partial\varphi(x^*).$$

So, problem $(GVIP)$ is a special case of the problem that consists in finding a zero of the sum of two operators. This last problem is considered, for example, in [18], [43], [57], [77], [93], [127].

A large variety of problems can be seen as special instances of problem $(GVIP)$. In the particular case where $\varphi$ is the indicator function of a nonempty closed convex subset $C$ of $H$, problem $(GVIP)$ reduces to the classical variational inequality problem:

$$(VIP) \begin{cases} \text{find } x^* \in C \text{ and } r(x^*) \in F(x^*) \text{ such that, for all } x \in C, \\ \langle r(x^*), x - x^* \rangle \geq 0. \end{cases}$$

This problem has been extensively studied in the literature. See, for example, [56], [102] for the singlevalued case, and [48], [61] for the multivalued case.

If $C$ is a closed convex cone of $H$ with $0 \in C$, then the problem reduces to

$$(GCP)\begin{cases} \text{find } x^* \in C \text{ and } r(x^*) \in F(x^*) \text{ such that,} \\ \\ r(x^*) \in C^* \text{ and } \langle r(x^*), x^* \rangle = 0, \end{cases}$$

where $C^*$ denotes the polar cone of $C$, defined by

$$C^* = \{c \in H : \langle c, u \rangle \geq 0, \forall u \in C\}.$$

This special case of problem $(GVIP)$ is known as a generalized complementarity problem. It was introduced by Karamardian (see [67]) and it is largely studied in the literature (see, for example, [62], [110]).

Finally, when $F$ is the subdifferential mapping of a finite-valued convex continuous function $f$ defined on $H$, problem $(GVIP)$ is just the nondifferentiable convex optimization problem:

$$(OP) \quad \min_{x \in H} \{f(x) + \varphi(x)\},$$

and problem $(VIP)$ reduces to the following constrained optimization problem:

$$(COP) \quad \min_{x \in C} f(x).$$

Algorithms that can be applied to solve problem $(GVIP)$ or one of its variants are too numerous to be enumerated here. Let us just mention some well-known and extensively studied classes of methods:

- proximal point algorithms (see [21], [24], [42], [43], [84], [109], [124]);
- splitting methods (see [29], [43], [51], [52], [57], [77], [101], [126], [127], [133]);
- auxiliary problem methods and cost approximation algorithms (see [33], [34], [40], [47], [99], [102], [105], [123], [126], [136]);
- interior point methods (see [94], [116], [129], [130], [131]);

- perturbation methods (see [45], [58], [82], [89], [125]).

The auxiliary problem principle was originally introduced by Cohen (see [30], [31]) and by Cohen and Zhu (see [34]) to solve problems such as the optimization problem $(OP)$. The general framework generated by this principle covers optimization algorithms ranging from gradient or subgradient methods to decomposition/coordination ones. In [33] and [85], this approach is applied to solve general variational inequalities such as problem $(GVIP)$. And more recently, this principle has been extended to nonsymmetric auxiliary operators by Renaud and Cohen (see [105]). The idea is to introduce a sequence of auxiliary operators $\{\Omega^k\}_{k\in I\!N}$ supposed to be strongly monotone and Lipschitz continuous but not necessarily symmetric, and positive numbers $\{\lambda_k\}_{k\in I\!N}$ so that $F$ be approximated at iteration $k$ by $\lambda_k^{-1}\Omega^k$. If $x^k$ denotes the current iterate at iteration $k$, the error made in approximating $F$ is taken into account by adding the error term $r(x^k) - \lambda_k^{-1}\Omega^k(x^k)$, where $r(x^k)$ denotes an element of $F(x^k)$. More precisely, the problem considered at iteration $k$ can be expressed as:

$$(AP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)), x - x^{k+1}\rangle \\ +\varphi(x) - \varphi(x^{k+1}) \geq 0. \end{cases}$$

Note that $r(x^k)$ can be given by a black box called at $x^k$.

In the first versions of the auxiliary problem method, each auxiliary operator $\Omega^k$ was chosen as the gradient of some continuously differentiable and strongly convex function $K^k$. In that case, subproblem $(AP^k)$ reduces to the following minimization problem:

$$(SAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \text{ the solution of} \\ \min_{x\in H}\{\, \lambda_k^{-1}K^k(x) + \varphi(x) + \langle r(x^k) - \lambda_k^{-1}\nabla K^k(x^k), x - x^k\rangle \,\}. \end{cases}$$

The variety of areas in which the auxiliary problem principle has brought about its contribution is very large. Let us just mention some of them, such as optimization (see [30], [31], [34]), saddle point problems (see [30], [31]), optimal control and dynamical systems (see [30]), stochastic optimization (see [39]), Nash equilibria (see [32]) and variational inequalities (see [33], [46], [47], [66], [85], [105]).

Many well–known algorithms for solving problem ($GVIP$) can be derived from the auxiliary problem scheme by choosing the auxiliary operators $\{\Omega^k\}_{k\in I\!N}$ (or $\{K^k\}_{k\in I\!N}$) in different ways. For example, we can recover linear approximation methods such as projection or Newton–like methods, or also proximal point methods, forward-backward methods,...

The convergence of some instances of the auxiliary problem scheme has already been studied in the literature. The case where $F$ is singlevalued is treated separately from the case where it is multivalued. When $F$ is *singlevalued*, the sequence of stepsizes $\{\lambda_k\}_{k\in I\!N}$ is supposed to be bounded away from zero and the convergence results are of two types. On the one hand, if the auxiliary operators are chosen symmetric, then $F$ is required to be strongly monotone (see, for example, [33], [46]) or to have the (pseudo) Dunn property (see, for example, [46], [85], [136]). On the other hand, if the auxiliary operators are not necessarily symmetric, then the operator $F$ and the sequence $\{\Omega^k\}_{k\in I\!N}$ have to be linked by a contraction condition (see [40], [99]) or by a kind of Dunn condition (see [105], [126], [136]). Now, when $F$ is *multivalued*, the sequence of stepsizes $\{\lambda_k\}_{k\in I\!N}$ converges to zero and the auxiliary operators are generally considered to be symmetric. In [33], Cohen proves the strong convergence of the scheme when $F$ is assumed to be strongly monotone. More recently, in [134], Zhu has obtained convergence results under weaker monotonicity assumptions. First, if $F$ is paramonotone and satisfies a continuity property, he shows that at least one weak limit point of the sequence generated by the algorithm is a solution of the original problem. Secondly, he proves weak convergence under a condition satisfied for example if the operator is paramonotone and compact-valued, or if it is the subdifferential of a lower semi-continuous proper convex function, or if it is strongly monotone.

In the case where the subproblems remain difficult to solve, several authors propose to approximate the function $\varphi$ by a sequence of more tractable functions. More precisely, at each iteration $k$, the original function $\varphi$ is replaced in the auxiliary subproblem $(AP^k)$ by an approximate function $\varphi^k$. Then the perturbed auxiliary subproblem can be expressed as:

$$(PAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)), x - x^{k+1}\rangle \\ +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

Approximations such as barrier functions, penalty functions, Tykhonov regularizations,... are encountered. The well-known variational convergence of Mosco (see [87]) encompasses various possibilities of data perturbations. In finite dimension, this notion coincides with that of epiconvergence (see, for example, [5], [7]). This kind of variational convergence notion is combined in the literature with various iterative methods such as proximal point algorithms (see [2], [13], [45], [64], [88], [89], [90], [91], [125]), splitting schemes (see [45], [57]) or Tykhonov algorithms (see [88], [125]).

For the auxiliary problem method, the convergence study of the perturbed scheme characterized by subproblems $(PAP^k)$ has already been initiated for some particular instances (see [73], [82], [120]). The most general result obtained for problem $(GVIP)$ (see [82]) is restricted to the case where $F$ is singlevalued and the auxiliary operators are symmetric. It ensures strong convergence provided that the sequence $\{\varphi^k\}_{k\in I\!N}$ converges to $\varphi$ in the sense of Mosco and the operator $F$ is strongly monotone and Lipschitz continuous. These conditions on $F$ are very restrictive and strongly limit the class of problems that can be considered. For example, problems with multiple solutions can't be treated. Nor can this result be applied to the case where $F = 0$. However, when $\varphi$ is not perturbed, convergence results with weaker conditions on $F$ as the Dunn property are available. Another failure of this result is to consider only symmetric auxiliary operators, what limits the application field. For example, the asymmetric projection method or the Newton method can't be analysed within the framework. On the other hand, to our knowledge, the influence of a variational perturbation of $\varphi$ on the convergence of the auxiliary problem scheme has not been studied in the literature when $F$ is multivalued.

Motivated by these shortcomings, our purpose in this work has been to obtain convergence results for the perturbed auxiliary problem scheme with iteration-dependent and nonsymmetric auxiliary operators, requiring conditions on $F$ as weak as possible such that we recover the same conditions on $F$ as in all the results mentioned above when $\varphi$ is not perturbed, as well in the singlevalued case as in the multivalued one. To achieve this, we have to work on the convergence properties of the sequence of approximate functions $\{\varphi^k\}_{k\in I\!N}$. Firstly, we assume that the sequence $\{\varphi^k\}_{k\in I\!N}$ approaches $\varphi$ in the sense of Mosco. Moreover, we focus our attention on interior approximations of the function $\varphi$, that is, on approximate functions $\varphi^k$ whose effective domains are contained in the domain of $\varphi$. In fact, we require that $\varphi \leq \varphi^k$ for all $k$. This allows us to consider, for example, barrier functions, interior approximations of the feasible set, Tykhonov approximations,... Secondly, in order to obtain the convergence of the scheme under weaker assumptions on $F$ than strong monotonicity, we have to require that the sequence $\{\varphi^k\}_{k\in I\!N}$ converges sufficiently fast to $\varphi$ by adding a condition on the speed of convergence of this sequence. For example, for the sequence of barrier functions, this condition amounts to a rate of convergence imposed on the barrier parameters. Note that this additional condition is needed both in the singlevalued case and in the multivalued one.

So, when $F$ is singlevalued, not only we prove the convergence for the general perturbed setting under weaker conditions on $F$ than strong monotonicity, but also we obtain conditions on $F$ that improve or reduce to the weakest ones existing in the nonperturbed case.

For the multivalued case, we also obtain convergence results under the same conditions on $F$ as in the best results when there is no perturbation. Moreover, we present a relaxation of this procedure by allowing an inexact computation of an element of $F(x^k)$. This is made by taking $r(x^k)$ in an enlargement of $F$ at $x^k$. This idea has already been used to relax other iterative methods in [25], [26], [118]. The $\epsilon$-enlargement notion used in these papers not being well suited for our purpose, we have introduced a new enlargement. So, we have been able to extend our convergence results to the inexact procedure. When $F$ is the subdifferential of a finite–valued convex continuous function $f$, the $\epsilon$–subdifferential can be chosen as the enlargement of $F = \partial f$ and our scheme reduces to the projected inexact subgradient procedure studied in [4].

Finally, we consider problem (GVIP) with $\varphi$ the sum of a lower semicontinuous proper convex function $p$ and the indicator function of a nonempty closed convex subset $C$ of $H$ such that $C \subseteq \text{dom } p \subseteq \text{dom } F$. When $F = 0$ and $C = H$, this problem reduces to minimize $p$ on $H$. In order to solve this optimization problem, some authors proposed the use of a bundle strategy (see, for example, [35]). Our aim has been to transpose this to our more general setting for solving approximately the auxiliary subproblems. However, to build a suitable piecewise linear convex approximation $p^k$ of $p$ as in the classical bundle strategy, the iterates need to remain in the interior of $C$. This is made possible by introducing a barrier function in the subproblems. So, we apply the arguments used in our previous results to prove the convergence of the resulting bundle scheme.

The thesis is organized as follows. The first chapter provides some basic definitions and results from the theory of convex analysis and nonlinear mappings related to our work. Some sufficient conditions for the existence of a solution of problem $(GVIP)$ are also recalled.

In the second chapter, we first illustrate the scope of the auxiliary problem procedure designed to solve problems like $(GVIP)$ by examining some well-known methods included in that framework. Then, we review the most representative convergence results for that class of methods that can be found in the literature in the case where $F$ is singlevalued as well as in the multivalued case. Finally, we somewhat discuss the particular case of projection methods to solve affine variational inequalities.

The third chapter introduces the variational convergence notion of Mosco and combines it with the auxiliary problem principle. Then, we recall the convergence conditions existing for the resulting perturbed scheme before our own contribution and we comment them. Finally, we introduce and illustrate the rate of convergence condition that we impose on the perturbations to obtain better convergence results.

Chapter 4 presents global and local convergence results for the family of perturbed methods in the case where $F$ is singlevalued. We also discuss how our results extend or improve the previous ones.

Chapter 5 studies the multivalued case. First, we present convergence results generalizing those obtained when there is no perturbations. Then, we relax the scheme by means of a notion of enlargement of an operator and we provide convergence conditions for this inexact scheme.

In Chapter 6, we build a bundle algorithm to solve problem $(GVIP)$ and we study its convergence.

Most results presented in this work have been published in international journals. The particular case of projection methods for affine variational inequalities is the subject of [55]. The results contained in Chapter 4 are a generalization in the infinite dimensional case of those appeared in [112]. Most results of Chapter 5 are the subject of [111] and [113]. The results of Chapter 6 are presented in [114].

# Chapter 1

# Background Notes

In this chapter, we recall some definitions and fundamental results related to the theory of convex analysis and of nonlinear mappings in a Hilbert space. We also recall some existence results for problem $(GVIP)$. We focus mainly on the background material needed to approach our work. The interested reader can find more comprehensive informations in these fields, for example, in [10], [38], [44], [68], [96], [100], [106], [128], [132].

Throughout this work, $H$ denotes a real Hilbert space equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the associated norm $\| \cdot \|$. If $x \in H$ and $\{x^k\}_{k \in \mathbb{N}}$ is a sequence of elements in $H$, $x^k \rightharpoonup x$ (resp. $x^k \to x$) expresses that the sequence $\{x^k\}_{k \in \mathbb{N}}$ weakly (resp. strongly) converges to $x$.
If $D$ is a symmetric positive definite matrix of $\mathbb{R}^n \times \mathbb{R}^n$, then the $D$–norm of a vector $x \in \mathbb{R}^n$ is given by

$$\|x\|_D = \sqrt{x^T D x},$$

and $\lambda_{min}(D)$, $\lambda_{max}(D)$ denote respectively, the minimum and maximum eigenvalues of $D$. For any square matrix $D$, $\mathrm{Ker}(D)$ and $\mathrm{Rank}(D)$ denote respectively its kernel and its rank, and $\mathrm{sym}(D)$ denotes its symmetric part, i.e.

$$\mathrm{sym}(D) = (D + D^T)/2.$$

If $C$ is a subset of $H$, $\mathrm{int}(C)$, $\mathrm{cl}(C)$ and $\mathrm{co}(C)$ denote respectively the interior, the closure and the convex hull of $C$.

## 1.1 Elements of Convex Analysis

Let $f : H \to I\!\!R \cup \{+\infty\}$ be a function. The effective domain of $f$ is the set

$$\mathrm{dom}\ f = \{x \in H : f(x) < +\infty\}.$$

The function $f$ is said to be proper if its effective domain is nonempty. We say that $f$ is convex if, for any $x, y \in H$ and $\lambda \in ]0,1[$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

If this inequality is strict whenever $x, y$ are different, the function $f$ is strictly convex. Moreover, $f$ is said to be strongly convex on $H$ if there exits a constant $\bar{\alpha} > 0$ such that, for any $x, y \in H$ and $\lambda \in ]0,1[$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - \bar{\alpha}\lambda(1-\lambda)\|x-y\|^2/2.$$

The function $f$ is lower semi-continuous on $H$ if, for each $x \in H$,

$$x^k \to x \Rightarrow \underline{\lim}_{k\to+\infty} f(x^k) \geq f(x).$$

When we consider the weak topology in $H$, the corresponding notion is the weak lower semi-continuity. Obviously, any weakly lower semi-continuous function is lower semi-continuous. The converse is not true in general but we have the following valuable property:

**Proposition 1.1** (See [44], Chapter I, Corollary 2.2) *Any convex function $f : H \to I\!\!R \cup \{+\infty\}$ is weakly lower semi-continuous if and only if it is lower semi-continuous.*

We denote by $\Gamma_0(H)$ the set of proper, convex, lower semi-continuous functions from $H$ into $I\!\!R \cup \{+\infty\}$.

We now introduce the notion of Gâteaux–differentiability.

**Definition 1.1** *Let $f : H \to I\!\!R \cup \{+\infty\}$. The directional derivative of $f$ at $x$ in the direction $d$ that we denote by $f'(x\,; d)$, is the limit as $\lambda \to 0^+$, if it exists, of*

$$\frac{f(x+\lambda d) - f(x)}{\lambda}. \tag{1.1}$$

*If there exists $s \in H$ such that $f'(x\,; d) = \langle d, s\rangle$ for all $d \in H$, then we say that $f$ is Gâteaux-(or G-)differentiable at $x$, we call $s$ the Gâteaux-(or G-)derivative of $f$ at $x$ and we denote it by $\nabla f(x)$.*

The uniqueness of the G-derivative follows directly. It is characterized by

$$\lim_{\lambda \to 0^+} \frac{f(x+\lambda d) - f(x)}{\lambda} = \langle d, \nabla f(x) \rangle, \; \forall d \in H.$$

If $f$ is convex, the expression (1.1) is an increasing function of $\lambda$ such that the limit always exists (but can be $-\infty$).

We next introduce the notion of subgradient of a (convex) function and we show how subdifferentiability constitutes a generalization of Gâteaux-differentiability.

**Definition 1.2** *Let $f : H \to I\!R \cup \{+\infty\}$. An element $s \in H$ is called a subgradient of $f$ at $x \in H$ if $f(x) \in I\!R$ and*

$$f(z) \geq f(x) + \langle s, z - x \rangle, \; \forall z \in H.$$

*The set of all subgradients of $f$ at $x$ is called the subdifferential of $f$ at $x$ and is denoted by $\partial f(x)$. If no subgradient exists at $x$, we say that $f$ is not subdifferentiable at $x$ and we set $\partial f(x) = \emptyset$.*

As an example, let us consider the indicator function of a nonempty closed convex subset $C$ of $H$, denoted by $\Psi_C$ and defined by

$$\Psi_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

By definition, $s \in \partial \Psi_C(x)$ if and only if $x \in C$ and

$$\Psi_C(z) \geq \Psi_C(x) + \langle s, z - x \rangle, \; \forall z \in H.$$

This means that $x \in C$ and $0 \geq \langle s, z - x \rangle$ for all $z \in C$ i.e., $s$ is normal to $C$ at $x$. Thus, $\partial \Psi_C(x)$ is the normal cone to $C$ at $x$:

$$\partial \Psi_C(x) = \begin{cases} \{s \in H : \langle s, z - x \rangle \leq 0, \; \forall z \in C\} & \text{if } x \in C \backslash \text{int}(C), \\ \{0\} & \text{if } x \in \text{int}(C), \\ \emptyset & \text{if } x \notin C. \end{cases}$$

The following proposition gives basic properties of the subdifferential of a lower semi–continuous and convex function.

**Proposition 1.2** (See [10], Chapter 4, Section 3, Theorem 17) *Let $f \in \Gamma_0(H)$. Then $f$ is subdifferentiable on int(dom $f$) and, for any $x \in$ int(dom $f$), $\partial f(x)$ is bounded, closed and convex.*

The following proposition shows that the subdifferential generalizes the G–derivative.

**Proposition 1.3** (See [44], Chapter I, Proposition 5.3) *Let $f : H \to I\!R \cup \{+\infty\}$ be a convex function. If $f$ is G–differentiable at $x \in H$, then $f$ is subdifferentiable at $x$ and $\partial f(x) = \{\nabla f(x)\}$. Conversely, if at a point $x \in H$, $f$ is continuous, finite and has only one subgradient, then $f$ is G–differentiable at $x$ and $\partial f(x) = \{\nabla f(x)\}$.*

Some interesting properties of the subdifferential of a strongly convex function are given in the following result.

**Proposition 1.4** *Let $f : H \to I\!R \cup \{+\infty\}$ be a subdifferentiable proper function. If $f$ is strongly convex with modulus $\bar{\alpha} > 0$, then for each $x, y \in H$ and each $r \in \partial f(x)$, $s \in \partial f(y)$, we have*

**(i)** $\langle r - s, x - y\rangle \geq \bar{\alpha}\|x - y\|^2$;

**(ii)** $f(y) \geq f(x) + \langle r, y - x\rangle + \bar{\alpha}\|x - y\|^2/2$.

**Proof.** This follows from the well-known fact that $f$ is strongly convex with modulus $\bar{\alpha}$ if and only if the function $f - \bar{\alpha}\|\cdot\|^2/2$ is convex (see [60], Chapter IV, Proposition 1.1.2). □

## 1.2 Preliminaries on Nonlinear Mappings

We first consider some basic notions for multivalued mappings. A multivalued mapping $T$ from $H$ to $H$ associates with any $x \in H$ a subset $T(x)$ of $H$, called the image or the value of $T$ at $x$. If, for each $x \in H$, the set $T(x)$ contains at most one element, then $T$ is said to be singlevalued.
The domain of the operator $T$ is the set

$$\text{dom } T = \{x \in H : T(x) \neq \emptyset\},$$

and the image of $T$ is the set

$$\text{Im } T = \cup_{x \in H} T(x) = \cup_{x \in \text{dom } T} T(x).$$

Actually, a multivalued mapping $T$ is characterized by its graph, the subset of $H \times H$ defined by

$$\text{Graph } T = \{(x, y) \in H \times H : y \in T(x)\}.$$

The inverse mapping $T^{-1}$ of $T$ is the multivalued map from $H$ to $H$ defined by

$$x \in T^{-1}(y) \Leftrightarrow y \in T(x).$$

Obviously, we have that dom $T^{-1} = \mathrm{Im}\, T$.
For given multivalued maps $T_1, T_2$ defined on $H$ and for fixed scalars $\alpha_1, \alpha_2 \in I\!R$, the mapping $\alpha_1 T_1 + \alpha_2 T_2$ is defined by

$$(\alpha_1 T_1 + \alpha_2 T_2)(x) = \begin{cases} \alpha_1 T_1(x) + \alpha_2 T_2(x) & \text{if } x \in \mathrm{dom}\ T_1 \cap \mathrm{dom}\ T_2, \\ \emptyset & \text{otherwise.} \end{cases}$$

The set of the operators on $H$ is ordered by the graph inclusion, i.e.

$$T_1 \subset T_2 \Leftrightarrow T_1(x) \subset T_2(x),\ \forall x \in H.$$

### 1.2.1 Continuity and Related Concepts

**Definition 1.3** *Let $T$ be a multivalued mapping defined on $H$. The operator $T$ is said to be upper semi-continuous at $x \in$ dom $T$, if to any neighborhood $V$ of the set $T(x)$, there corresponds a neighborhood $U$ of $x$ such that $T(U) \subset V$.*
*$T$ is said to be weakly–strongly (resp. strongly–weakly) continuous at $x \in$ dom $T$ if it is upper semi-continuous at $x$ from the weak (resp. strong) topology on $H$ into the strong (resp. weak) topology on $H$.*
*$T$ is said to be bounded if it carries bounded subsets of dom $T$ into bounded subsets of $H$.*
*$T$ is said to be weakly–strongly (resp. strongly–weakly) closed if it follows from $\{x^k\}_{k \in I\!N} \subset$ dom $T$, $x^k \rightharpoonup x$ (resp. $x^k \to x$), $r^k \in T(x^k), r^k \to r$ (resp. $r^k \rightharpoonup r$), that $r \in T(x)$. And $T$ is said to be weakly closed if it follows from $\{x^k\}_{k \in I\!N} \subset$ dom $T$, $x^k \rightharpoonup x$, $r^k \in T(x^k), r^k \rightharpoonup r$, that $r \in T(x)$.*
*We say that $T$ is upper hemi-continuous at $x \in$ dom $T$ if (dom $T$ is convex and) for any $y \in$ dom $T$, the mapping defined on $[0,1]$ by:*

$$\lambda \to \{ \langle r_\lambda, y - x \rangle : r_\lambda \in T(x + \lambda(y - x)) \}$$

*is upper semi-continuous at $0^+$.*
*$T$ is said to be Lipschitz continuous on a subset $B$ of $H$ if*

$$\exists\, L > 0 \text{ such that } \forall x, y \in B \quad e(T(x), T(y)) \leq L\, \|x - y\|,$$

*where $e(T(x), T(y)) = \sup_{r \in T(x)} \inf_{s \in T(y)} \|r - s\|$.*

The next lemma will be used in the sequel.

**Lemma 1.1** *Let $B$ be a bounded subset of $H$. If $T$ is Lipschitz continuous on $B$ and if there exists $\bar{y} \in B$ such that $T(\bar{y})$ is bounded, then $T$ is bounded on $B$, i.e., there exists $c > 0$ such that $\|r(x)\| \leq c$ for all $x \in B$ and $r(x) \in T(x)$.*

**Proof.** Let $\epsilon > 0$. Then, by assumption, $e(T(x), T(\bar{y})) \leq L\|x - \bar{y}\|$ for all $x \in B$, i.e.,

$\forall x \in B, \forall r(x) \in T(x),\ \exists\, r(\bar{y}) \in T(\bar{y})$ such that $\|r(x) - r(\bar{y})\| \leq L\|x - \bar{y}\| + \epsilon$.

Since $B$ and $T(\bar{y})$ are bounded, there exist $c_1 > 0$ and $c_2 > 0$ such that $\|x\| \leq c_1$ for all $x \in B$ and $\|r(\bar{y})\| \leq c_2$ for all $r(\bar{y}) \in T(\bar{y})$. Then, for all $x \in B$ and $r(x) \in T(x)$, we have successively

$$\begin{aligned} \|r(x)\| &\leq \|\, r(x) - r(\bar{y}) \,\| + \|\, r(\bar{y}) \,\| \\ &\leq L\,[\, \|x\| + \|\, \bar{y} \,\| \,] + \epsilon + c_2 \\ &\leq L\,[\, c_1 + \|\, \bar{y} \,\| \,] + \epsilon + c_2, \end{aligned}$$

i.e., what we have to prove. □

In the case of a singlevalued mapping, we use the following terminology.

**Definition 1.4** *Let $T$ be singlevalued. We say that $T$ is continuous (resp. weakly continuous) if for any sequence $\{x^k\}_{k \in I\!N} \subset$ dom $T$, $x^k \to x$ (resp. $x^k \rightharpoonup x$), we have that $T(x^k) \to T(x)$ (resp. $T(x^k) \rightharpoonup T(x)$).*
*$T$ is said to be weakly–strongly (resp. strongly–weakly) continuous, if $\{x^k\}_{k \in I\!N} \subset$ dom $T$, $x^k \rightharpoonup x$ (resp. $x^k \to x$) implies that $T(x^k) \to T(x)$ (resp. $T(x^k) \rightharpoonup T(x)$).*
*$T$ is said to be compact if it is continuous and for any bounded subset $B$ of $H$, the image $T(B)$ is relatively compact.*
*$T$ is hemi-continuous at $x \in$ dom $T$ if (dom $T$ is convex and) for any $y \in$ dom $T$, the map $t \to T(x + t(y - x))$ is continuous from $[0, 1]$ into the weak topology of $H$.*
*$T$ is Lipschitz continuous with constant $L > 0$ if for any $x, y \in H$,*

$$\|T(x) - T(y)\| \leq L\|x - y\|.$$

*$T$ is nonexpansive if it is Lipschitz continuous with constant $L = 1$.*

The concept of coercivity plays an important role to ensure that an operator be onto. It is defined here after under different forms.

**Definition 1.5** *The mapping $T$ is coercive on $H$ if*

$$\lim_{\|x\|\to+\infty} \frac{1}{\|x\|} \inf_{r\in T(x)} |\langle r, x\rangle| = +\infty.$$

*It is weakly coercive on $H$ if*

$$\lim_{\|x\|\to+\infty} \inf_{r\in T(x)} \|r\| = +\infty.$$

*It is coercive with respect to the element $h \in H$ if there exist a number $\rho > 0$ and an element $x_0 \in$ dom $T$ such that $x \in$ dom $T$ and $\|x\| \geq \rho$ imply that*

$$\langle r, x - x_0\rangle > \langle h, x - x_0\rangle, \, \forall r \in T(x).$$

Since $\langle r, x\rangle \leq \|r\| \, \|x\|$, we have that coercivity implies weak coercivity. It is also easy to see that if $T$ is coercive on $H$, it will be coercive with respect to any element $h$ of $H$ (See [100], Chapter III, Section 2.8). Moreover, the coercivity condition is satisfied whenever the mapping is defined on a bounded domain.

### 1.2.2 Derivative of a Singlevalued Mapping

**Definition 1.6** *A singlevalued mapping $T : D \subset H \to H$ is Gâteaux-(or G-)differentiable at an interior point $x$ of $D$ if there exists a linear operator $A : H \to H$ such that, for any $d \in H$,*

$$\lim_{\lambda\to 0} \frac{\|T(x + \lambda d) - T(x) - \lambda A d\|}{\lambda} = 0. \tag{1.2}$$

*There exists at most one linear operator $A$ for which (1.2) is satisfied, it is denoted by $\nabla T(x)$ and is called the Gâteaux-(or G-)derivative of $T$ at $x$.*

So, if $T$ is G-differentiable at $x \in H$, then, for any fixed $d \in H$, the mapping $G(\lambda) = T(x + \lambda d)$ is differentiable at zero and

$$\nabla G(0) = \lim_{\lambda\to 0} \frac{G(\lambda) - G(0)}{\lambda} = \nabla T(x)\, d.$$

Consequently, $G$ is continuous at zero and the following proposition follows:

**Proposition 1.5** (See [96], Chapter 3, Point 3.1.4. and NR 3.1-2) *If $T : D \subset H \to H$ is G-differentiable at $x \in D$, then $T$ is hemi-continuous at $x$.*

The following mean-value theorem will be used in our analysis.

**Proposition 1.6** (See [96], Chapter 3, Point 3.2.5. and NR 3.2-4) *If $T : D \subset H \to H$ is G-differentiable on the convex set $D_0 \subset D$, then for any $x, y, z \in D_0$,*

$$\|T(y) - T(z) - \nabla T(x)(y - z)\| \leq \sup_{0 \leq t \leq 1} \|\nabla T(z + t(y - z)) - \nabla T(x)\| \, \|y - z\|.$$

**Definition 1.7** *A mapping $T : D \subset H \to H$ is said to be continuously G–differentiable on the convex set $D_0 \subset D$ if $T$ is G–differentiable at each $x \in D_0$ and $\nabla T$, the G-derivative of $T$, is continuous on $D_0$.*

### 1.2.3 Monotonicity and Related Topics

**Definition 1.8** *Let $C$ be a closed convex subset of $H$. The mapping $T$ is monotone on $C$ if, for any $x, y \in C$ and any $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x) - r(y), x - y \rangle \geq 0.$$

If $T_1, T_2$ are monotone, then $T_1^{-1}, \lambda T_1 (\lambda \geq 0)$, $T_1 + T_2$ are monotone. One important subclass of monotone operators is given in the following proposition.

**Proposition 1.7** (See [10], Chapter 4, Section 3, Proposition 9) *If $f : H \to I\!R \cup \{+\infty\}$ is a proper convex function, then its subdifferential mapping $\partial f$ is monotone.*

The properties of the resolvent operator corresponding to a monotone operator are given in the following proposition.

**Proposition 1.8** (See [10], Chapter 6, Section 6, Proposition 8) *If $T$ is monotone on $H$, then for each $\lambda > 0$, the resolvent operator $(I + \lambda T)^{-1}$ is a singlevalued nonexpansive map from Im $(I + \lambda T)$ to $H$.*

The following result gives a characterization of monotonicity for differentiable mappings.

**Proposition 1.9** (See [63], Proposition 4.1) *Assume that $T$ is G–differentiable on $H$ and $C$ is a nonempty closed convex subset of $H$. Then $T$ is monotone on $C$ if and only if sym$(\nabla T(x))$ is positive semi-definite for each $x \in C$.*

A basic property of monotone mappings is local boundedness.

**Definition 1.9** *A multivalued mapping $T$ is locally bounded at $x \in H$ if there exists a neighbourhood $U$ of $x$ such that the set*

$$T(U) = \cup_{y \in U \cap dom\ T} T(y)$$

*is bounded in $H$. Moreover, $T$ is locally bounded on the set $D$ if $T$ is locally bounded at each point of $D$.*

**Proposition 1.10** (See [100], Chapter III, Section 2.2) *A monotone mapping defined on $H$ is locally bounded at the interior points of dom $T$.*

In general, any strongly–weakly continuous and singlevalued operator is hemi-continuous. The converse is also true for monotone operators.

**Proposition 1.11** (See [100], Chapter III, Section 2.2) *Any monotone hemi-continuous and singlevalued operator $T$ is strongly-weakly continuous on int(dom $T$).*

We can introduce the concept of maximal monotonicity.

**Definition 1.10** *A monotone mapping $T$ is maximal if there is no other monotone operator whose graph strictly contains the graph of $T$.*

It follows that $T$ is maximal monotone if and only if it is monotone and for each $x, y \in H$,

$$\langle x - \xi, y - \eta \rangle \geq 0,\ \forall (\xi, \eta) \in \text{Graph } T \Rightarrow y \in T(x).$$

The following characterization is crucial in the study of maximal monotone mappings.

**Proposition 1.12** (See [18], Chapitre 2, Proposition 2.2) *Let $T$ be a mapping defined on $H$. Then the following assertions are equivalent:*

**(i)** *$T$ is maximal monotone,*

**(ii)** *$T$ is monotone and Im $(I + T) = H$,*

**(iii)** *for each $\lambda > 0$, the resolvent operator $(I + \lambda T)^{-1}$ is nonexpansive and dom $((I + \lambda T)^{-1}) = H$.*

Note that if $T$ is maximal monotone on $H$, then $T^{-1}$ and $\lambda T$ (with $\lambda > 0$) are also maximal monotone. Moreover, a very important subclass of maximal monotone operators is given in the following proposition.

**Proposition 1.13** (See [100], Chapter III, Section 2.13) *Let $f \in \Gamma_0(H)$. Then the subdifferential operator $\partial f$ is maximal monotone.*

Fundamental properties of maximal monotone operators are displayed in the following proposition.

**Proposition 1.14** (See [10], Chapter 6, Section 7, Proposition 3) *Let $T$ be a maximal monotone operator. Then for any $x \in$ dom $T$, $T(x)$ is closed and convex. Moreover, $T$ is weakly–strongly (or strongly–weakly) closed.*

In the singlevalued case, the following proposition can be interesting to characterize maximality of a monotone operator.

**Proposition 1.15** (See [18], Chapitre 2, Corollaire 2.5) *Le $T$ be a monotone singlevalued operator such that dom $T = H$. The following properties are equivalent:*

**(i)** *$T$ is maximal monotone,*

**(ii)** *$T$ is strongly–weakly closed,*

**(iii)** *$T$ is strongly–weakly continuous,*

**(iv)** *$T$ is hemi-continuous.*

We also recall a criterion for maximal monotone operators to be onto.

**Proposition 1.16** (See [132], Chapter 32, Corollary 32.35) *Let $T$ be a maximal monotone operator. Assume that one of the two following conditions holds:*

**(i)** *dom $T$ is bounded,*

**(ii)** *$T$ is weakly coercive.*

*Then $T$ is onto.*

For two maximal monotone operators $T_1$ and $T_2$, we have that $T_1 + T_2$ is monotone, but it does not necessarily follow that $T_1 + T_2$ is maximal monotone. Indeed, for example, we can have that dom $T_1 \cap$ dom $T_2 = \emptyset$. The following results give additional conditions to require.

**Proposition 1.17** (See [108], Theorem 1) *Let $T_1, T_2$ be maximal monotone operators defined on $H$. Suppose that either one of the following conditions is satisfied:*

**(i)** *int(dom* $T_1$*)*$\cap$ *dom* $T_2 \neq \emptyset$,

**(ii)** *there exists an element* $x \in$ *cl(dom* $T_1$*)*$\cap$ *cl(dom* $T_2$*) such that* $T_1$ *is locally bounded at* $x$.

*Then* $T_1 + T_2$ *is maximal monotone.*

**Proposition 1.18** (See [18], Chapitre 2, Lemme 2.4) *If* $T_1$ *is a maximal monotone operator defined on* $H$ *and* $T_2$ *is a Lipschitz continuous and monotone operator from* $H$ *to* $H$, *then* $T_1 + T_2$ *is maximal monotone.*

**Proposition 1.19** (See [108], Theorem 3) *Let* $T_1$ *be the subdifferential operator of the indicator function of a nonempty closed convex subset* $C$ *of* $H$ *and* $T_2$ *be a singlevalued monotone operator such that* $C \subset$ *dom* $T_2$ *and* $T_2$ *is hemi-continuous on* $C$, *then* $T_1 + T_2$ *is a maximal monotone operator.*

Some results for maximal monotone operators can be weakened by using the following concept of pseudomonotonicity.

**Definition 1.11** *Let* $C$ *be a closed convex subset of* $H$. *The multivalued mapping* $T$ *is said to be pseudomonotone in the sense of Brézis ([17]) on* $C$ *if for each sequence* $\{x^k\}_{k\in I\!N} \subset C$, *it follows from* $r^k \in T(x^k)$, $x^k \rightharpoonup x$ *and* $\overline{\lim}\langle r^k, x^k - x\rangle \leq 0$, *that for each* $y \in C$, *there corresponds an element* $r_y \in T(x)$ *such that*

$$\langle r_y, x - y\rangle \leq \underline{\lim}_{k\to+\infty}\langle r^k, x^k - y\rangle.$$

**Proposition 1.20** (See [132], Chapter 32, Problem 32.3a) *Any singlevalued weakly–strongly continuous mapping on* $C$ *is pseudomonotone in the sense of Brézis on* $C$.

**Proposition 1.21** (See [132], Problem 32.3c) *Assume that* $T$ *is a multivalued and monotone operator defined on* $C$. *If* $T$ *is strongly–weakly continuous on line segments of* $C$ *and for each* $x \in C$, $T(x)$ *is nonempty, closed and convex, then* $T$ *is pseudomonotone in the sense of Brézis on* $C$.

**Proposition 1.22** (See [100], Chapter III, Section 2.4) *Any maximal monotone operator defined on* $H$ *and such that dom* $T = H$ *is pseudomonotone in the sense of Brézis on* $H$.

Stronger monotonicity conditions are defined here below.

**Definition 1.12** *Let $C$ be a closed convex subset of $H$ and $T$ be a multivalued mapping defined on $H$. $T$ is said to be strictly monotone on $C$ if it is monotone on $C$ and for all $x, y \in C$, and all $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x) - r(y), x - y \rangle = 0 \Rightarrow x = y.$$

*It is strongly monotone on $C$ if there exists a positive constant $\bar{\alpha}$ such that, for each $x, y \in C$, $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x) - r(y), x - y \rangle \geq \bar{\alpha} \|x - y\|^2.$$

Other generalized monotonicity concepts are also used in the literature.

**Definition 1.13** *Let $\varphi \in \Gamma_0(H)$ and let $C$ be a closed convex subset of dom $\varphi$. $T$ is said to be pseudomonotone over $C$ if for all $x, y \in C$, and all $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x), y - x \rangle \geq 0 \Rightarrow \langle r(y), y - x \rangle \geq 0.$$

*$T$ is $\varphi$-pseudomonotone over $C$ if for all $x, y \in C$, and all $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x), y - x \rangle + \varphi(y) - \varphi(x) \geq 0 \Rightarrow \langle r(y), y - x \rangle + \varphi(y) - \varphi(x) \geq 0.$$

Note that the pseudomonotonicity concept of Definition 1.11 is not to be confused with the concepts presented in Definition 1.13.

### 1.2.4 Dunn Property

The Dunn property is a concept of generalized monotonicity for singlevalued mappings that lies strictly between simple and strict monotonicity. This property was introduced by Browder and Petryshyn in [22] in the context of computing fixed point solutions. It has been later used in [23], [51] to establish the convergence of the projection algorithm or also in [85], [46], [105] in the framework of the auxiliary problem principle. Several other names are used in the literature as "cocoercivity" ([83], [126], [135], [136]), "strong–f–monotonicity" ([80]) and "firm–nonexpansiveness" ([43]).

**Definition 1.14** *Let $C$ be a closed convex subset of $H$. A singlevalued operator $T$ is said to have the Dunn property over $C$ with modulus $\gamma > 0$ if for all $x, y \in C$,*

$$\langle T(x) - T(y), x - y \rangle \geq \gamma \|T(x) - T(y)\|^2.$$

The following weaker forms will be also used in this work.

**Definition 1.15** *Let $\varphi \in \Gamma_0(H)$, and let $C$ be a closed convex subset of dom $\varphi$. $T$ has the pseudo Dunn property over $C$ with modulus $\gamma > 0$ if for all $x, y \in C$,*

$$\text{if } \langle T(x), y - x\rangle \geq 0 \text{ holds, then}$$

$$\langle T(y), y - x\rangle \geq \gamma \|T(x) - T(y)\|^2.$$

*$T$ has the $\varphi$–pseudo Dunn property over $C$ with modulus $\gamma > 0$ when for all $x, y \in C$,*

$$\text{if } \langle T(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0 \text{ holds, then}$$

$$\langle T(y), y - x\rangle + \varphi(y) - \varphi(x) \geq \gamma \|T(x) - T(y)\|^2.$$

It is straightforward that if $T$ has the Dunn property over $C$, then it has the $\varphi$–pseudo Dunn property over $C$. The converse is not true in general as it can be seen by taking $T(x) = 1/x$, $C = [1, +\infty[$ and $\varphi = 0$. It is obvious that the Dunn property is strictly weaker than strict or strong monotonicity (consider, for example, the constant mapping). The Dunn property can also be stated as the strong monotonicity of the inverse mapping $T^{-1}$ with constant $\gamma > 0$. When this property holds, the operator $T$ is monotone and Lipschitz continuous with constant $\gamma^{-1}$. Thus, the fact that $T$ is singlevalued is implied by the Dunn property itself. Moreover, in the case where $T$ is the G–derivative of a convex function, we have the following result:

**Proposition 1.23** (See [14], Corollaire 10) *Let $C$ be a closed convex subset of $H$. If $f : H \to I\!R$ is convex and G–differentiable, then the two following assertions are equivalent:*

**(i)** *$\nabla f$ is Lipschitz continuous with constant $L > 0$ on $C$,*

**(ii)** *$\nabla f$ has the Dunn property with constant $L^{-1}$ on $C$.*

However, this equivalence is not true in the general nonsymmetric case. For example, if for all $x_1, x_2 \in I\!R$, $T(x_1, x_2) = (-x_2, x_1)$, then $T$ is Lipschitz continuous and monotone on $I\!R^2$ but does not enjoy the Dunn property. Nevertheless, we have the following result:

**Proposition 1.24** *If $T$ is strongly monotone with constant $\bar{\alpha} > 0$ over $C$ and Lipschitz continuous over $C$ with constant $L > \bar{\alpha}$, then $T$ enjoys the Dunn property with constant $\bar{\alpha}/L^2$ over $C$.*

**Proof.** This is straightforward from the definitions. $\square$

In the case where $T$ is G–differentiable, we can define the following differentiable form of the Dunn property:

**Definition 1.16** *Let $C$ be a closed convex subset of $H$. If the mapping $T$ is G–differentiable, we say that it satisfies the differentiable form of the Dunn property on $C$ if there exists some constant $\gamma > 0$ such that the matrix $\nabla T(x)^T - \gamma \nabla T(x)^T \nabla T(y)$ is positive semi-definite for all $x, y \in C$. If this matrix is positive semi-definite for all $x = y \in C$, then we say that $T$ satisfies the weak differentiable form of the Dunn property.*

It is shown in [80] that the differentiable form of the Dunn property implies the Dunn property and that the converse is true on any open convex subset of $C$. Sufficient conditions imposed on $\nabla T$ and $C$ to ensure these differentiable forms are discussed in [80].

### 1.2.5 Paramonotonicity

The notion of paramonotonicity is a concept that lies strictly between monotonicity and strict monotonicity. In the case of a singlevalued mapping, it is also strictly weaker than the Dunn property. It was first used in [23] where it was given no name, then introduced in [28] and further studied in [63]. The name "monotonicity–plus" is also used in the literature (See, for example, [37], [134]).

**Definition 1.17** *A multivalued operator $T$ is said to be paramonotone on a convex subset $C$ of $H$ if it is monotone on $C$ and for all $x, y \in C$, and $r(x) \in T(x), r(y) \in T(y)$,*

$$\langle r(x) - r(y), x - y \rangle = 0 \Rightarrow r(y) \in T(x) \text{ and } r(x) \in T(y).$$

If $T$ is singlevalued, this condition amounts to require that for all $x, y \in C$,

$$\langle T(x) - T(y), x - y \rangle = 0 \Rightarrow T(x) = T(y).$$

According to the definitions, we have that strict monotonicity on $C$ implies paramonotonicity on $C$ which in turn implies monotonicity on $C$. It is easy to see that these classes are distinct even in the case of linear mappings. For example, if $T(x_1, x_2) = (x_1 - x_2, x_1)$, for all $x_1, x_2 \in \mathbb{R}$, then $T$ is monotone but not paramonotone; if $T(x_1, x_2) = (x_1, 0)$, for all $x_1, x_2 \in \mathbb{R}$, then $T$ is

paramonotone but not strictly monotone.
The next proposition shows that paramonotonicity encompasses at least the convex optimization case.

**Proposition 1.25** (See [24], Proposition 1(i)) *If $T$ is the subdifferential of a convex function $f : H \to I\!R$, then $T$ is paramonotone on $H$.*

The relevance of paramonotonicity appears in the convergence analysis of algorithms which generate a sequence $\{x^k\}_{k\in I\!N}$ expected to converge to a solution of the problem $(GVIP)$. Indeed, in many cases one can verify only that a weak limit point $\bar{x}$ of $\{x^k\}_{k\in I\!N}$ satisfies the optimality inequality for some solution $x^*$. When the operator is paramonotone, this will be sufficient to conclude that $\bar{x}$ is also a solution. This feature is proved in the next proposition which generalizes Proposition 2.3 of [63].

**Proposition 1.26** *Let us consider problem $(GVIP)$ with $\varphi \in \Gamma_0(H)$. Assume that $F$ is paramonotone on dom $\varphi$, and let $x^*$ be a solution of problem $(GVIP)$. Then $\bar{x}$ solves problem $(GVIP)$ if there exists $\bar{r} \in F(\bar{x})$ such that*

$$\langle \bar{r}, x^* - \bar{x}\rangle + \varphi(x^*) - \varphi(\bar{x}) \geq 0.$$

**Proof.** Assume that

$$\langle \bar{r}, x^* - \bar{x}\rangle + \varphi(x^*) - \varphi(\bar{x}) \geq 0, \tag{1.3}$$

for some $\bar{r} \in F(\bar{x})$ and some solution $x^*$ of problem $(GVIP)$ (with $r(x^*) \in F(x^*)$ associated with $x^*$). Since $F$ is monotone and $(x^*, r(x^*))$ is a solution of problem $(GVIP)$, we have that

$$\langle \bar{r}, \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) \geq \langle r(x^*), \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) \geq 0. \tag{1.4}$$

Combining (1.3) and (1.4), we obtain that

$$\begin{aligned} 0 &= \langle \bar{r}, \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) \\ &= \langle r(x^*), \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*). \end{aligned} \tag{1.5}$$

And thus, $\langle \bar{r} - r(x^*), \bar{x} - x^*\rangle = 0$, which implies that $r(x^*) \in F(\bar{x})$ since $F$ is paramonotone. We deduce also from (1.5) that, for all $x \in$ dom $\varphi$:

$$\langle r(x^*), x - \bar{x}\rangle + \varphi(x) - \varphi(\bar{x}) \geq 0.$$

So, we conclude that $\bar{x}$ is a solution of problem $(GVIP)$ and the proof is complete. □

Finally, let us mention that when $\varphi$ is the indicator function of a closed convex subset of $H$, paramonotonicity is used to ensure convergence of several interior point methods with generalized distances (see, for example, [24], [25], [28]).

In the singlevalued case, the definition of paramonotonicity reveals that it is equivalent to strict monotonicity of the inverse mapping. So, the Dunn property, which is equivalent to strong monotonicity of the inverse mapping, is a property stronger than paramonotonicity. For example, the mapping defined by $T(x) = x^3$, for all $x \in I\!R$, is paramonotone but does not satisfy the Dunn property on $I\!R$. When $T$ is the G–derivative of a convex function $f$, Proposition 1.23 amounts to say that $T$ satisfies the Dunn property if and only if $T$ is paramonotone and Lipschitz continuous. However, it is not true in general. Indeed, if we consider for all $x_1, x_2 \in [1, +\infty[$, $T(x_1, x_2) = (\sqrt{x_1} + x_2, \sqrt{x_2} - x_1)$, then $T$ is paramonotone and Lipschitz continuous on $[1, +\infty[\times[1, +\infty[$ but it does not satisfy the Dunn property on this set.

Consider now that $T$ is singlevalued and continuously differentiable. Sufficient conditions for paramonotonicity are given in the following proposition:

**Proposition 1.27** (See [63], Proposition 4.2 or [37], Proposition 9) *Assume that $T$ is continuously differentiable on a closed convex subset $C$ of $H$ with nonempty interior. If for any $x \in C$, $sym(\nabla T(x))$ is positive semi-definite and one of the three following statements holds:*

**(i)** $\langle \nabla T(x)\, h, h\rangle = 0 \Rightarrow \nabla T(x)\, h = 0,\ \forall h \in H,$

**(ii)** $Ker(sym(\nabla T(x))) \subset Ker(\nabla T(x)),$

**(iii)** $Rank(\nabla T(x)) \leq Rank(sym(\nabla T(x))),$

*then $T$ is paramonotone on $C$.*

This result provides easily checkable conditions since in order to check that a monotone and differentiable operator is paramonotone, it suffices to verify that its Jacobian matrix does not loose rank when it is symmetrized. Let us point out that conditions (ii) and (iii) are trivially satisfied when $\nabla T(x)$ is symmetric for all $x \in C$ i.e., when $T$ is the G–derivative of a function defined on $C$.

### 1.2.6 The Case of Affine Operators in $I\!R^n$

In this section, we consider operators of the form $T(x) = Ax + b$ with $A \in I\!R^{n\times n}, b \in I\!R^n$. Let $C$ be a closed convex subset of $I\!R^n$ with nonempty interior. According to the definitions of generalized monotonicity, we have that $T$ is monotone on $C$ if and only if $A$ is positive semi-definite on $C$, and that $T$ is strictly monotone on $C$ if $A$ is positive definite on $C$ (See [63], Proposition 3.2). In a natural way, we say that $A$ has the Dunn property with modulus $\gamma > 0$ on $C$ if it is the case for $T$ i.e.,

$$\langle Ax, x\rangle \geq \gamma \|Ax\|^2, \; \forall x \in C.$$

To describe the subclass of paramonotone operators, we need to introduce the class of positive semi-definite–plus matices.

**Definition 1.18** *The matrix $A \in I\!R^{n\times n}$ is said to be positive semi-definite–plus (psd–plus) if $A$ is positive semi-definite and for any $h \in I\!R^n$,*

$$\langle Ah, h\rangle = 0 \Rightarrow Ah = 0.$$

It is well–known that any symmetric positive semi-definite matrix is psd–plus (this follows from the fact that $A$ can be written as $A = B^T B$ for some $B \in I\!R^{n\times n}$ by operating a Choleski factorization of $A$). It can also be easily seen that if $A$ and $A^2$ are both positive semi-definite, then $A$ is psd–plus (See [37], Proposition 10). However, the converse of the above assertions does not hold. So, the following proposition provides necessary and sufficient conditions for a positive semi-definite matrix to be psd–plus.

**Proposition 1.28** (See [37], Proposition 11) *Let $A$ be an $(n \times n)$ positive semi-definite matrix. Then the three following conditions are equivalent:*

**(i)** *$A$ is psd–plus,*

**(ii)** *$Ker(sym(A)) \subset Ker(A)$,*

**(iii)** *$Rank(A) \leq Rank(sym(A))$.*

For more details about psd–plus matrices, we refer to [55], [78] and [80].

Obviously, $T$ is paramonotone on $C$ if and only if $A$ is psd–plus (See [37], Proposition 9). Moreover, it has been shown in [135] (Proposition 3.4) that $A$ has the Dunn property on $C$ if and only if $A$ is psd–plus. Thus, in the affine case, paramonotonicity and the Dunn property coincide. To conclude, let us sum up our comments in one result:

**Proposition 1.29** *Let $C \subset I\!R^n$ be a convex set with nonempty interior and take $T(x) = Ax + b$. Then*

**(i)** *$T$ is monotone on $C$ if and only if sym$(A)$ is positive semi-definite,*

**(ii)** *$T$ is paramonotone on $C$*
*if and only if $T$ has the Dunn property on $C$*
*if and only if sym$(A)$ is positive semi-definite and Ker(sym$(A)) \subset$ Ker$(A)$,*

**(iii)** *$T$ is strictly monotone on $C$*
*if and only if sym$(A)$ is positive semi-definite and Ker(sym$(A)) = \{0\}$ i.e., sym$(A)$ is positive definite.*

## 1.3 Existence Theory

As observed in the introduction, the general variational inequality problem $(GVIP)$ amounts to find a zero of the operator $T$ which is the sum of $F$ and $\partial\varphi$. One basic result ensuring existence of a zero of a maximal monotone operator is the following:

**Theorem 1.1** (See [108], Proposition 2) *Let $T$ be a maximal monotone operator on $H$. Suppose that there exists $\alpha > 0$ such that*

$$\langle x, y \rangle \geq 0, \ \textit{whenever} \ \|x\| > \alpha, x \in \textit{dom } T, y \in T(x).$$

*Then there exists an element $x^* \in H$ such that $0 \in T(x^*)$.*

Let us give some comments on the main condition of this theorem. It obviously holds, for example, if the effective domain of $T$ is a bounded set or if $T$ is coercive on $H$. When $T$ is the sum of two operators $T_1$ and $T_2$, $T$ is coercive for example if $0 \in$ dom $T_1$ and $T_2$ is coercive (or vice versa) or if dom $T_1 \cap$ dom $T_2$ is bounded.

Recall that in our case, $T = F + \partial\varphi$ and the subdifferential mapping of a proper closed convex function is known to be maximal monotone. The application of Theorem 1.1 gives the following result:

**Theorem 1.2** (See [132], Proposition 32.36) *Assume that the following assumptions are satisfied:*

- *$F$ is monotone on $H$;*

- *$\varphi \in \Gamma_0(H)$;*
- *one of the following three conditions holds:*

  **(i)** *$F$ is singlevalued and hemi-continuous,*

  **(ii)** *$F$ is maximal monotone and int(dom $F$)$\cap$ dom $\partial\varphi \neq \emptyset$,*

  **(iii)** *$F$ is maximal monotone and dom $F \cap$ int(dom $\partial\varphi) \neq \emptyset$;*

- *The sum $F + \partial\varphi$ is coercive with respect to 0 i.e., there exist $\rho > 0$ and $x_0 \in$ dom $F \cap$ dom $\partial\varphi$ such that*

$$\langle r, x - x_0 \rangle > 0, \forall (x, r) \in \textit{Graph } (F + \partial\varphi) \textit{ with } \|x\| > \rho.$$

*Then problem $(GVIP)$ admits at least one solution.*

When $F$ is not maximal monotone but is only pseudomonotone in the sense of Brézis, we dispose of the following existence result for problem $(GVIP)$:

**Theorem 1.3** (See [132], Problem 32.4, Theorem 32.A) *If the following assumptions are satisfied:*

- *$F$ is pseudomonotone in the sense of Brézis and bounded;*
- *$\varphi \in \Gamma_0(H)$ and dom $\varphi \subset$ dom $F$;*
- *For each $x \in$ dom $\varphi$, the set $F(x)$ is a nonempty closed convex subset of $H$;*
- *$F$ is strongly-weakly continuous on simplices i.e., for each finite subset $S \subset$ dom $\varphi$, the map $F : co(S) \to 2^H$ is strongly–weakly continuous;*
- *$F$ is $\partial\varphi$–coercive with respect to 0 i.e., there are $x_0 \in$ dom $\partial\varphi$ and $\rho > 0$ such that $\langle r, x - x_0 \rangle > 0$ for all $(x, r) \in$Graph $F$ with $\|x\| > \rho$.*

*Then problem $(GVIP)$ has a solution.*

Other existence results for problem $(GVIP)$ (or $(VIP)$) under pseudomonotonicity or other generalized monotonicity assumptions can be found, for example, in [10], [36], [41], [48], [70], [72], [117]. In these papers, $F$ is assumed to be upper semi-continuous or upper hemi-continuous.

Concurrently to this, it should be noticed that many real live problems lead to equations governed by a noncoercive operator. A general approach

for studying the solvability of such noncoercive problems relies on the asymptotic behavior of the sets, functions or operators involved in the problem. It is called the recession approach.
Consider the problem of finding a zero of a maximal monotone operator $T$. The recession function associated to $T$ is the support function of the closed convex set cl(Im $T$):

$$f_\infty^T(x) = \sup_{y \in \mathrm{cl}(\mathrm{Im}\ T)} \langle x, y \rangle.$$

The next theorem ensures the existence of a zero of a non necessarily coercive operator $T$.

**Theorem 1.4** (See [6], Theorem 4.1) *Let $T$ be a maximal monotone operator defined on $H$. If the two following conditions are satisfied:*

**(i)** *compactness condition:*

$$\forall t_k \to +\infty, \forall x^k \rightharpoonup x \text{ with } T(t_k x^k) \text{ bounded, we have that } x^k \to x;$$

**(ii)** *compatibility condition:*

$$f_\infty^T \geq 0 \text{ and } \ker f_\infty^T = \{x \in H : f_\infty^T(x) = 0\} \text{ is a subspace },$$

*then there exists at least one element $x^* \in H$ such that $0 \in T(x^*)$.*

To apply this theorem when $T$ is the sum of two maximal monotone operators $T_1$ and $T_2$, it should be interesting to know when $f_\infty^{T_1+T_2} = f_\infty^{T_1} + f_\infty^{T_2}$. For that purpose, let us first introduce the following condition: $T$ is said to satisfy the Brézis–Haraux condition (See [19]) if

$$\sup_{(m,n) \in \mathrm{Graph}\ T} \langle n - y, x - m \rangle < +\infty, \ \forall x \in \ \mathrm{dom}\ T, \forall y \in \ \mathrm{Im}\ T.$$

Situations where a given operator satisfies the Brézis–Haraux condition are given in [19]. For example, this condition holds for the subdifferential of a lower semi-continuous proper convex function. And the result is the following:

**Proposition 1.30** (See [6]) *Let $T_1$ and $T_2$ be two maximal monotone operators defined on $H$. If $T_1$ and $T_2$ satisfy the Brézis–Haraux condition and cl$(T_1 + T_2)$ is maximal monotone, then $f_\infty^{T_1+T_2} = f_\infty^{T_1} + f_\infty^{T_2}$.*

For further results based on the recession approach concerning noncoercive variational inequalities, we refer to [1], [6], [53].

# Chapter 2

# The Auxiliary Problem Method

In this chapter, we first introduce the auxiliary problem framework designed to solve problems like $(GVIP)$ and we illustrate its scope by providing some examples of well–known methods that can be obtained as particular cases. Then we recall the most representative convergence results that can be found in the literature for the general auxiliary problem scheme or one instance of it.

## 2.1 The Auxiliary Problem Framework and Particular Instances

As presented in the introduction, the auxiliary problem principle generates, at iteration $k$, the following subproblem:

$$(AP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)), x - x^{k+1}\rangle \\ +\varphi(x) - \varphi(x^{k+1}) \geq 0, \end{cases}$$

where $\{\Omega^k\}_{k\in I\!N}$ is a sequence of auxiliary operators supposed to be strongly monotone and Lipschitz continuous and $\{\lambda_k\}_{k\in I\!N}$ is a sequence of positive numbers.

Subproblem $(AP^k)$ can also be equivalently written under the following inclusion form:

$$(AP^k)\begin{cases} \text{find } x^{k+1} \in H \text{ such that,} \\ 0 \in F(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)) + \partial\varphi(x^{k+1}). \end{cases}$$

The assumptions imposed on $\Omega^k$ and $\varphi$ ensure that each subproblem $(AP^k)$ admits one and only one solution. Indeed, since $\partial\varphi$ is maximal monotone, $\lambda_k$ is a positive real number, $\Omega^k$ is monotone and Lipschitz continuous, we have that $\lambda_k^{-1}\Omega^k + \partial\varphi$ is maximal monotone (see Proposition 1.18). Moreover, $\lambda_k^{-1}\Omega^k + \partial\varphi$ is strongly monotone and thus coercive. It follows then from Theorem 1.1 that subproblem $(AP^k)$ admits at least one solution. Moreover, strict monotonicity of $\lambda_k^{-1}\Omega^k + \partial\varphi$ implies that this solution is unique.

Recall that when the auxiliary operator $\Omega^k$ is chosen as the gradient of some continuously differentiable and strongly convex function $K^k$, problem $(AP^k)$ reduces to the minimization problem:

$$(SAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \text{ the solution of} \\ \min_{x\in H}\{\, \lambda_k^{-1}K^k(x) + \varphi(x) + \langle r(x^k) - \lambda_k^{-1}\nabla K^k(x^k), x - x^k\rangle \,\}. \end{cases}$$

The fact that $F$ appears only through the linear part $\langle r(x^k), x-x^k\rangle$ in the minimization subproblem $(SAP^k)$ has motivated the use of this method to build up decomposition algorithms. Parallel decomposition can be achieved when $H$ is the product of $N$ Hilbert spaces and $\varphi$ is additive with respect to this decomposition. Indeed, problem $(SAP^k)$ splits up into $N$ independent minimization subproblems provided that the auxiliary function $K^k$ is chosen additive with respect to the structure. Combination of this approach with the relaxation principle leads to relaxed algorithms studied in [30], [31], [34] for differentiable and nondifferentiable optimization. Besides the optimization field, the auxiliary problem principle has been applied in many areas (see the references cited in the introduction).

Strongly related to the class of auxiliary problem methods is the cost approximation framework of Patriksson (see [102]). In that work, the solution of the auxiliary problem $(AP^k)$ is used to define a feasible search direction. Then, a step is taken in that direction through a (possibly inexact) line search with respect to a merit function for problem $(GVIP)$. This step defines a new iteration point and the process is repeated. Note that Patriksson limits himself to the case where the operator $F$ is singlevalued and $H$ is of finite dimension.

In the sequel, we will see that by appropriately choosing the sequence $\{\Omega^k\}_{k\in I\!N}$ (or $\{K^k\}_{k\in I\!N}$), many well–known algorithms to solve problem $(GVIP)$ fall within the auxiliary problem scheme. The fact that the auxiliary operator can change at each iteration allows a great degree of flexibility and is crucial for application to methods like Newton's method for example.

- **LINEAR APPROXIMATION METHODS**

  Consider the classical variational inequality problem $(VIP)$ in $H = I\!R^n$ and let the sequence of auxiliary operators $\{\Omega^k\}_{k\in I\!N}$ take the form

  $$\Omega^k(x) = D(x^k)\,x,\ \forall k \in I\!N,\ \forall x \in I\!R^n,$$

  where $D(x^k)$ is a $(n \times n)$ positive definite matrix. So, with $\lambda_k = 1$ for all $k$, subproblem $(AP^k)$ reduces to

  $$\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and find } x^{k+1} \in C \text{ such that,} \\ \langle r(x^k) + D(x^k)(x^{k+1} - x^k), x - x^{k+1}\rangle \geq 0,\ \forall x \in C. \end{cases} \tag{2.1}$$

  Adequate choices of the matrix $D(x^k)$ lead to different well–known methods:

  - **Projection Methods**

    If we take $D(x^k) = G^k$, with $G^k$ a symmetric positive definite matrix, then it is easy to see that the point $x^{k+1}$ solving subproblem (2.1) is precisely the projection of the point $x^k - (G^k)^{-1}r(x^k)$ onto the closed convex set $C$ with respect to the $G^k$–norm, i.e.

    $$\begin{cases} x^{k+1} = \mathrm{proj}_C^{G^k}(x^k - (G^k)^{-1}r(x^k)) \\ \text{with } r(x^k) \in F(x^k), \end{cases} \tag{2.2}$$

where for a given vector $z$, $\|z\|_{G^k} = \sqrt{z^T G^k z}$ and $\mathrm{proj}_C^{G^k} z$ is the unique solution of

$$\min_{x \in C} \|z - x\|_{G^k}.$$

Projection methods for $(VIP)$ are studied, for example, in [55], [83], [99], [126].

When $F$ is the subdifferential mapping of a function $f$ and $G^k = I$, subproblem (2.2) characterizes the projected subgradient method to solve the constrained optimization problem $(COP)$. We refer to [4], [34], and the references cited therein for more details on this procedure.

– **Newton–like Methods**

When $F$ is singlevalued and continuously differentiable, the following choices for $D(x^k)$ are possible:

$$\begin{array}{lll} D(x^k) & = \nabla F(x^k) & \text{(Newton)}, \\ & \approx \nabla F(x^k) & \text{(quasi–Newton)}, \\ & = \mathrm{sym}(\nabla F(x^k)) & \text{(symmetrized Newton)}, \\ & = \mathrm{diag}(\nabla F(x^k)) & \text{(linearized Jacobi)}, \\ & = \frac{\mathrm{l}(\nabla F(x^k)) + \mathrm{diag}(\nabla F(x^k))}{\omega} & \text{(successive over-relaxation)}, \\ & = \frac{\mathrm{u}(\nabla F(x^k)) + \mathrm{diag}(\nabla F(x^k))}{\omega}, & \end{array}$$

with

$$\begin{array}{ll} \mathrm{diag}(\nabla F(x^k)) & = \text{the diagonal part of } \nabla F(x^k), \\ \mathrm{l}(\nabla F(x^k)) & = \text{the lower triangular part of } \nabla F(x^k), \\ \mathrm{u}(\nabla F(x^k)) & = \text{the upper triangular part of } \nabla F(x^k), \\ \omega \in ]0, 2[. & \end{array}$$

We refer to [99] for more details and references about these methods.

Obviously, when $F$ is the gradient of a function $f$ twice differentiable, we recover the classical Newton–type methods for solving the constrained optimization problem $(COP)$.

- **SPLITTING METHODS**

Consider problem $(GVIP)$ under the inclusion form:

$$\text{find } x^* \in H : 0 \in F(x^*) + \partial\varphi(x^*).$$

The auxiliary problem framework can be viewed as a splitting scheme since it exploits the special structure of problem $(GVIP)$ by treating separately $F$ and $\partial\varphi$. Indeed, subproblem $(AP^k)$ can be equivalently written under a forward–backward form where a forward step for $F$ is followed by a backward step for $\partial\varphi$:

$$(AP^k) \begin{cases} x^{k+1} = (\Omega^k + \lambda_k \partial\varphi)^{-1}(\Omega^k(x^k) - \lambda_k r(x^k)), \\ \\ \text{with } r(x^k) \in F(x^k). \end{cases}$$

If we take for $\Omega^k$ the identity mapping, then we recover the classical forward–backward scheme. This algorithm has been extensively studied in the literature (see, for example, [29], [51], [77], [101], [127]).

Other classes of methods can be derived by choosing an auxiliary operator $\Omega^k$ which depends on $F$ or some part of $F$. For example, if we choose $\Omega^k(x) = x + \lambda_k F(x)$ (assuming that $F$ is singlevalued), the forward–backward procedure reduces to a backward step and thus turns into the proximal point algorithm for the mapping $F + \partial\varphi$:

$$x^{k+1} = (I + \lambda_k(F + \partial\varphi))^{-1} x^k.$$

The literature about the proximal point method and its generalizations is very large. We just mention here some representative papers such as [43], [84], [109].

These two applications follow from extreme choices for the mapping $\Omega^k$. Now, consider the splitting of $F$, $F = F_1 + F_2$ with $F_1$ singlevalued. We can make an intermediate choice for the auxiliary operator $\Omega^k$ by taking:

$$\Omega^k(x) = x + \lambda_k F_1(x).$$

This choice leads to the following procedure:

$$\begin{cases} x^{k+1} = (I + \lambda_k(F_1 + \partial\varphi))^{-1}(x^k - \lambda_k r_2(x^k)), \\ \\ \text{with } r_2(x^k) \in F_2(x^k). \end{cases}$$

Obviously, the extreme splitting where $F_1 = 0$ and $F_2 = F$ leads to the typical forward–backward scheme while the splitting where $F_1 = F$ and $F_2 = 0$ reduces to the proximal point procedure.

Each algorithm that we have shown to be a particular instance of the auxiliary problem method includes a large variety of methods depending on the problem formulation and the context of application. For further references about these methods, we refer to the book of Patriksson (see [102]) where he shows the connections among classes of algorithms for the solution of instances of problem $(GVIP)$ and he provides a large number of examples that can be described within this framework.

## 2.2 Convergence of the Auxiliary Problem Method

An important research work is devoted to the convergence study of the auxiliary problem type algorithms. In recent years, the convergence has been established under weaker and weaker monotonicity assumptions on $F$. In the sequel, we recall the most representative results existing for the scheme described by subproblems $(AP^k)$ or one particular instance. We will treat separately the case where $F$ is singlevalued and the case where it is multivalued. Indeed, the multivalued case requires in general stronger assumptions on $F$. Moreover, the selection rule for the stepsizes $\{\lambda_k\}_{k\in\mathbb{N}}$ is different. In the singlevalued case, the sequence $\{\lambda_k\}_{k\in\mathbb{N}}$ is bounded away from zero while in the multivalued case, this sequence converges (but not too fast) to zero. We find the same phenomenon when we pass from differentiable to nondifferentiable optimization (compare [31] and [34]).

Before going on further, let us mention that if the strong monotonicity of the operator $F$ is in general superfluous to ensure convergence of the auxiliary problem method, monotonicity of $F$ without any additional condition does not suffice. To illustrate this, consider the problem of finding a zero of the operator $F$ of rotation by $\pi/2$ in $\mathbb{R}^2$. So, $F$ is the linear and monotone mapping defined by $F(x_1, x_2) = (-x_2, x_1)$, for all $x_1, x_2 \in \mathbb{R}$. If we choose $\Omega^k = I$ and $\lambda_k = \lambda$, for each $k \in \mathbb{N}$, subproblem $(AP^k)$ reduces to

$$x^{k+1} = x^k - \lambda F(x^k), \text{ i.e.}$$

$$\begin{cases} x_1^{k+1} = x_1^k + \lambda x_2^k, \\ x_2^{k+1} = x_2^k - \lambda x_1^k. \end{cases}$$

Hence, we have

$$\|x^{k+1}\|^2 = (1 + \lambda^2)\,\|x^k\|^2.$$

So that, the norm of the iterates strictly increases for any positive $\lambda$. And thus, the sequence generated can't converge towards the unique zero $x^* = (0, 0)$ of the operator $F$.

### 2.2.1 The Singlevalued Case

When $F$ is singlevalued, the convergence results can be classified in two groups following that the conditions imposed on $F$ concern $F$ alone or $F$ and the sequence $\{\Omega^k\}_{k\in I\!N}$ together. In the first group, the auxiliary operators are symmetric and assumptions like -$F$ is strongly monotone- (see, for example, [33], [46]) or -$F$ has the (pseudo) Dunn property- (see, for example, [46], [85], [136]) are considered. In the second group, the auxiliary operators are not necessarily symmetric and assumptions linking $F$ and $\{\Omega^k\}_{k\in I\!N}$ can take the form of a contraction condition as in [40], [99] or the form of a Dunn condition as in [105], [126] and [136].

For the first group, the next result establishes the convergence of the algorithm defined by the minimization subproblems $(SAP^k)$ when the auxiliary function is not iteration dependent so that $K^k = K$ for all $k$.

**Theorem 2.1** *Let $H = I\!R^n$ and $\varphi \in \Gamma_0(H)$. Assume that the solution set of problem $(GVIP)$ is nonempty, $K : H \to I\!R$ is continuously differentiable and strongly convex over dom $\varphi$ (with modulus $\beta > 0$), and its derivative $\nabla K$ is Lipschitz continuous over dom $\varphi$ (with constant $\Lambda > 0$). If $F$ is continuous and satisfies the Dunn property over dom $\varphi$ with modulus $\gamma > 0$ and there exists $\underline{\lambda} > 0$ such that $\underline{\lambda} < \lambda_k < 2\beta\gamma$ for all $k$, then the sequence $\{x^k\}_{k\in I\!N}$ generated by solving subproblems $(SAP^k)$ with $K^k = K$ for all $k$, is bounded and converges to some solution of problem $(GVIP)$.*

The proof of this result can be found in [136] (Theorem 3.2) when $\lambda_k = \lambda$ for all $k$ and $\varphi = f + \Psi_C$ with $f$ a continuous and finite–valued convex function and $C$ a closed convex subset of $I\!R^n$, or in [46] (Theorem 3 and Corollary 1) when $\varphi = \Psi_C$ and the Dunn condition on $F$ is replaced by the weaker requirement that $F$ satisfies the pseudo Dunn condition over $C$.

When $F$ is strongly monotone over dom $\varphi$ with constant $\bar{\alpha} > 0$ and Lipschitz continuous over dom $\varphi$ with constant $L > 0$, then it obviously satisfies the (pseudo) Dunn condition with modulus $\bar{\alpha}/L^2$ (see Proposition 1.24). Hence the condition "$\underline{\lambda} < \lambda_k < 2\beta\gamma$, for all $k$" reduces to "$\underline{\lambda} < \lambda_k < 2\beta\bar{\alpha}/L^2$, for all $k$". In that case, Theorem 2.1 can be directly transposed to an infinite dimensional space $H$ with as conclusion that the sequence $\{x^k\}_{k\in I\!N}$ strongly converges to the unique solution of problem $(GVIP)$. By this way, we recover the convergence results of [33] (Theorem 2.2) and [46] (Theorem 2).

Now, if we allow the auxiliary operator to be non necessarily symmetric, the convergence results at our disposal impose a condition that links $F$ and $\Omega^k$. In order to solve the classical variational inequality problem $(VIP)$ in the finite dimensional case, Pang and Chan consider in [99], the linearization methods described by problems $(AP^k)$ with $\Omega^k(x) = D(x^k)\,x$ where $D(x^k)$ is an $(n \times n)$ matrix and $\lambda_k = 1$ for all $k$ (see subproblem (2.1)). They obtain the following result:

**Theorem 2.2** (See [99], Theorem 2.9) *Let $H = I\!R^n$ and let $C$ be a nonempty closed convex subset of $H$. Assume that problem $(VIP)$ admits a solution, $F$ is continuous on $C$, $D(\cdot)$ is bounded on bounded subsets of $C$, and the following condition holds:*

**Condition (PCH):**
*there exists a symmetric positive definite matrix $G$ such that, for all $x \in C$, $D(x) - G$ is positive semi-definite on $C$ and there exists a positive constant $b < 1$ such that, for all $x, y \in C$:*

$$\|G^{-1}[F(y) - F(x) - D(y)(y - x)]\|_G \leq b\|y - x\|_G.$$

*Then the sequence $\{x^k\}_{k\in I\!N}$ generated by the linear approximation method (2.1) converges to a solution of problem $(VIP)$ for any initial vector $x^0 \in C$.*

Note that Pang and Chan also provide a local version of this result and apply it to methods like Newton.

This contraction approach is also used by Dafermos in [40] to solve problem $(VIP)$ in finite dimension. To see that the auxiliary problems $(AP^k)$ can be encompassed in the general framework of Dafermos, let us write the auxiliary operators under the form $\Omega(x^k, \cdot)$ instead of $\Omega^k$. With this convention and if $\lambda_k = 1$ for all $k \in I\!N$, subproblem $(AP^k)$ can be seen as a particular instance of Dafermos scheme:

$$\begin{cases} \text{find } x^{k+1} \in C \text{ such that, for all } x \in C, \\ \langle g(x^{k+1}, x^k), x - x^{k+1} \rangle \geq 0, \end{cases}$$

by choosing:

$$g(x, y) = F(y) + \Omega(y, x) - \Omega(y, y).$$

Then the following result can be deduced:

**Theorem 2.3** (See [40], Theorem 2.1) *Let $H = I\!R^n$ and let $C$ be any nonempty closed convex subset of $H$. Assume that $C$ is bounded, $F$ is continuously differentiable, the auxiliary operator $\Omega$ is differentiable in $x$ and $y$ and has a positive definite Jacobian with respect to $y$ in $H \times H$. Moreover, suppose that the following norm condition holds:*

**Condition (DAF):**

$$\|[sym(\Omega_y(y_1, x_1))]^{-1/2} \left[\nabla F(y_2) + \Omega_x(y_2, x_2) - \Omega_x(y_2, y_2) - \Omega_y(y_2, y_2)\right]$$

$$[sym(\Omega_y(y_3, x_3))]^{-1/2}\| < 1, \quad \forall x_1, x_2, x_3, y_1, y_2, y_3 \in C,$$

*where the subscripts $x$ and $y$ indicate partial differentiation.*

*Then the sequence $\{x^k\}_{k \in I\!N}$ generated by solving problems $(AP^k)$ with $\lambda_k = 1$ for all $k \in I\!N$, converges to the unique solution of problem $(VIP)$.*

Dafermos shows that Condition $(DAF)$ implies that $F$ is stricly monotone on $C$ (see [40], Proposition 2.3) such that problem $(VIP)$ has at most one solution. Moreover, as observed by Dafermos, the requirement that $C$ is bounded can be dropped if $\Omega$ is strongly monotone with respect to $y$ and if the left-hand side of Condition $(DAF)$ remains bounded away from 1 uniformly in $x_1, x_2, x_3, y_1, y_2, y_3 \in C$ i.e.

**Condition (SDAF):**
*there exists a positive constant $b < 1$ such that,*

$$\|[sym(\Omega_y(y_1, x_1))]^{-1/2} \left[\nabla F(y_2) + \Omega_x(y_2, x_2) - \Omega_x(y_2, y_2) - \Omega_y(y_2, y_2)\right]$$

$$[sym(\Omega_y(y_3, x_3))]^{-1/2}\| < b, \quad \forall x_1, x_2, x_3, y_1, y_2, y_3 \in C.$$

But in that case, it can be proven that $F$ is strongly monotone on $C$ (see [79], Section 4, Proposition 4).

The Conditions $(DAF)$ and $(SDAF)$ are difficult to verify when $\Omega$ is iteration dependent since they involve any points $x_1, x_2, x_3, y_1, y_2, y_3$ in $C$. In their study of averaging schemes, Magnanti and Perakis (see [79]) consider a similar condition but with $x_1 = x_2 = x_3 = y_1 = y_2 = y_3$. In that case, Condition $(SDAF)$ appears as a differential form of the following one:

**Condition ($\overline{\mathbf{PCH}}$):**
*there exists a symmetric positive definite matrix $G$ such that, for all $x \in C$, $\Omega(x, \cdot) - G$ is monotone over $C$ and there exists a positive constant $b < 1$ such that, for all $x, y \in C$:*

$$\|G^{-1}[F(y) - F(x) - (\Omega(y, y) - \Omega(y, x))]\|_G \leq b\|y - x\|_G.$$

Observe that this last condition reduces to Condition $(PCH)$ when $\Omega(x, y) = D(x)\,y$.
However, Dafermos provides examples showing that Condition $(DAF)$ corresponds to the classical convergence condition when applied, for example, to the projection algorithm.

Other results are based on a kind of Dunn condition that links $F$ and $\{\Omega^k\}_{k \in I\!N}$. For example, we can use the result of Renaud and Cohen (see [105]) to derive the convergence of the auxiliary problem method defined by problems $(AP^k)$ with $\lambda_k = \lambda$ for all $k$ and

$$\Omega^k = \Omega = \nabla h + \lambda L, \; \forall k \in I\!N, \tag{2.3}$$

where $h$ is a continuously differentiable function from $H$ to $I\!R$ and $L$ is a singlevalued mapping from $H$ into $H$. Note that this decomposition of $\Omega$ is quite natural as it will be explained in Chapter 4. Then we obtain the following result:

**Theorem 2.4** (See [105], Theorem 3.4) *Let $\varphi \in \Gamma_0(H)$. Assume that problem $(GVIP)$ admits at least one solution and that $F + \partial\varphi$ is maximal monotone. Let the derivative of $h$ be strongly monotone with modulus $\beta > 0$ on $\operatorname{dom}\varphi$ and Lipschitz continuous over any bounded subset of $H$, and let $L - F$ be hemi-continuous. Moreover, suppose that the following condition holds:*

**Condition (RC):**
*There exists a positive constant* $\gamma$ *such that, for all* $x, y \in$ *dom* $\varphi$*, for all* $\varphi_x \in \partial\varphi(x)$ *and* $\varphi_y \in \partial\varphi(y)$*:*

$$\langle F(x) + \varphi_x - F(y) - \varphi_y, x - y\rangle \geq \gamma\|F(x) - L(x) - F(y) + L(y)\|^2.$$

*If* $0 < \lambda < 2\beta\gamma$*, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *generated by solving subproblems* $(AP^k)$ *with* $\Omega^k$ *taking the form (2.3), is bounded, and every weak limit point is a solution of problem* $(GVIP)$*. Moreover, if* $\nabla K$ *is weakly continuous on* $H$*, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *weakly converges to a solution of problem* $(GVIP)$*. Finally, if* $F + \partial\varphi$ *is strongly monotone, then* $\{x^k\}_{k\in\mathbb{N}}$ *strongly converges to the unique solution of* $(GVIP)$*.*

A slightly stronger condition for convergence is imposed in [136] (Theorem 4.1) for the case where $H$ is a finite dimensional space. Indeed, instead of $(RC)$, they require $L$ to be monotone and $F - L$ to satisfy the Dunn property. Tseng proposes also a similar condition (see [126], Proposition 2) for the more particular case where $\Omega^k(x) = Dx$ for all $k$, where $D$ is an $(n \times n)$ positive definite matrix which is not necessarily symmetric. Moreover, he shows that this condition is implied by Condition $(PCH)$ when $F$ is Lipschitz continuous (see [126], Proposition 3).

### 2.2.2 The Multivalued Case

In the multivalued case, the auxiliary operators are generally given by gradient mappings such that $\Omega^k = \nabla K^k$ for all $k$. Moreover, the stepsizes $\{\lambda_k\}_{k\in\mathbb{N}}$ are chosen such that

$$\lambda_k > 0, \forall k \in \mathbb{N}, \sum_{k=0}^{+\infty} \lambda_k^2 < +\infty, \text{ and } \sum_{k=0}^{+\infty} \lambda_k = +\infty. \tag{2.4}$$

As it can be seen in the convergence proofs, this selection rule ensures that the stepsizes are small enough to guarantee boundedness of the sequence but not too small to ensure convergence to a solution of the problem. This rule is also considered in the literature for non smooth minimization problems (see [4], [12], [34], [35], [103]). When $F$ is supposed to be strongly monotone, Cohen proves the strong convergence of the auxiliary problem scheme $(SAP^k)$ to the unique solution of problem $(GVIP)$.

**Theorem 2.5** (See [33], Theorem 3.1) *Let $\varphi \in \Gamma_0(H)$. Assume that problem $(GVIP)$ admits at least one solution, $K$ is differentiable and strongly convex over dom $\varphi$ (with modulus $\beta > 0$). If the sequence $\{\lambda_k\}_{k\in I\!N}$ satisfies the rule (2.4), $F$ is strongly monotone over dom $\varphi$ (with modulus $\bar{\alpha}$) and is such that*

$$\exists\, a, b > 0 : \, \|r(x)\| \leq a\|x\| + b, \forall x \in dom\ \varphi, \forall r(x) \in F(x),$$

*then the sequence $\{x^k\}_{k\in I\!N}$ generated by solving subproblems $(SAP^k)$ with $K^k = K$ for all $k$, strongly converges to the unique solution of problem $(GVIP)$.*

Motivated by the fact that, in many applications, the mapping $F$ is not strongly monotone, Zhu proposed very recently in [134] a result based on the concept of paramonotonicity which lies between monotonicity and strong monotonicity. So, Zhu obtains the following result for the auxiliary problem scheme $(SAP^k)$ to solve problem $(VIP)$.

**Theorem 2.6** (See [134], Theorem 1) *Let $C$ be any nonempty closed convex subset of $H$. Assume that problem $(VIP)$ admits at least one solution and that the following conditions are satisfied:*

**(i)** *for each $k \in I\!N, K^k : H \to I\!R$ is continuously differentiable and strongly convex with modulus $\beta_k \geq \beta > 0$ over $C$;*

**(ii)** *for each $k \in I\!N$, there exists a positive number $\eta_k$ such that, for all $x, y \in C$:*

$$K^{k+1}(x) - K^{k+1}(y) - \langle \nabla K^{k+1}(y), x - y\rangle$$
$$\leq \eta_k(K^k(x) - K^k(y) - \langle \nabla K^k(y), x - y\rangle);$$

**(iii)** *there exists a positive number $\beta'$ such that, for all $k \in I\!N$,*

$$\beta_k/M^k \geq \beta',$$

*where $M^k = \prod_{j=0}^{k-1} \eta_j$;*

**(iv)** *the sequence $\{\lambda_k/M^k\}_{k\in I\!N}$ satisfies the rule (2.4);*

**(v)** *$F$ is paramotone over $C$;*

**(vi)** *for any bounded sequence* $\{z^k\}_{k\in\mathbb{N}}$ *of* $C$*, the sequence* $\{r^k\}_{k\in\mathbb{N}}$ *with* $r^k \in F(z^k)$ *is bounded, and there exist subsequences* $\{z^{k'}\}_{k'\in K\subset\mathbb{N}}$ *and* $\{r^{k'}\}_{k'\in K\subset\mathbb{N}}$ *such that*

$$z^{k'} \rightharpoonup \bar{z},\ r^{k'} \rightharpoonup \bar{r}, \bar{r} \in F(\bar{z}),\ \text{and} \lim_{k'\to\infty} \langle r^{k'}, z^{k'}\rangle \geq \langle \bar{r}, \bar{z}\rangle;$$

**(vii)**

$$\exists\, a, b > 0 : \|r(x)\| \leq a\|x\| + b,\ \forall x \in C, \forall r(x) \in F(x).$$

*Then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *generated by solving the auxiliary subproblems* $(SAP^k)$ *with* $\varphi = \Psi_C$ *is bounded and at least one of its weak limit points is a solution of problem* $(VIP)$*.*

Moreover, Zhu shows that any weak limit point is a solution if in addition to the assumptions of Theorem 2.6 (unless condition (vi)), the operator $F$ satisfies the following property (where $x^*$ denotes a solution of problem $(VIP)$):

**Condition** $(ZH_{x^*})$**:**
*There exist a constant* $\alpha > 0$ *and a function* $l$ *Lipschitz continuous on any compact subset of* $C$ *such that,*

$$\begin{aligned}
&\langle r(y), y - x^*\rangle \geq \alpha(l(y) - l(x^*)),\ \forall y \in C,\ \forall r(y) \in F(y),\ \text{and}\\
&l(y) = l(x^*),\ \text{if } y \text{ is a solution of problem } (VIP),\\
&l(y) > l(x^*),\ \text{otherwise.}
\end{aligned}$$

As observed by Zhu, this property holds for some important mappings such as strongly monotone multivalued mappings, or paramonotone and Lipschitz continuous mappings, or the subdifferential mappings of any functions of $\Gamma_0(H)$ which are Lipschitz continuous on any compact subset of $C$. More details about this kind of condition will be given in Chapter 5.

## 2.3 Projection Methods for Singular and Linear Variational Inequalities in Finite Dimension

In this section, we give some comments on the results presented in [55] by Goeleven, Stavroulakis, Salmon and Panagiotopoulos. In that paper, the authors study projection methods to solve affine variational inequalities of the form:

$$(AVIP)\begin{cases} \text{find } x^* \in C \text{ such that, for all } x \in C,\\ \langle Mx^* - q, x - x^*\rangle \geq 0, \end{cases}$$

where $C$ is a nonempty closed convex subset of $I\!R^n$, $M$ is an $(n \times n)$ positive semi-definite matrix and $q$ is a vector of $I\!R^n$. This is a particular instance of problem $(VIP)$ with $F(x) = Mx - q$ and $H = I\!R^n$. This type of problem is encountered, for example, in unilateral contact applications where $x$ and $x^*$ are displacement vectors, $C$ is the kinematically admissible set, $q$ is the loading initial strain vector and $M$ is the stiffness matrix. As discussed in [55], if rigid body displacements or rotations can't be excluded, then the resulting matrix $M$ is singular. This motivates the development of algorithms which do not require the matrix $M$ to be positive definite. The purpose of paper [55] is to show that symmetric and asymmetric projection methods are particularly well suited to solve this kind of problem and to show how theory applies to concrete engineering problems as the unilateral cantilever problem or the elastic stamp problem.

For the symmetric projection method, recall that the iterates are generated by the following rule:

$$x^{k+1} = (G + \lambda \partial \Psi_C)^{-1}(Gx^k - \lambda(Mx^k - q)),$$

where $G$ is a symmetric positive definite matrix and $\lambda > 0$. The convergence of this sequence $\{x^k\}_{k \in I\!N}$ to a solution of problem $(AVIP)$ is ensured (see [55], Theorem 2.2) provided that problem $(AVIP)$ admits at least one solution, that the matrix $M$ has the Dunn property with modulus $\gamma > 0$ and $0 < \lambda < 2\gamma / \|G^{-1}\|$. Note that this result follows directly from Theorem 2.1. Now, when $M$ is only positive semi-definite but does not have the Dunn property, Goeleven et al. propose the use of an asymmetric projection method described as follows:

$$x^{k+1} = ((G - \lambda Q) + \lambda \partial \Psi_C)^{-1}((G - \lambda Q)x^k - \lambda(Mx^k - q)),$$

where $G$ is a symmetric positive definite matrix, $\lambda > 0$, and $Q$ is the skew–symmetric matrix defined by $Q = (M^T - M)/2$. It is shown in [55] that if the problem $(AVIP)$ admits at least one solution and $0 < \lambda < 4\|M + M^T\|^{-1} / \|G^{-1}\|$, then the sequence generated by this asymmetric projection method converges to a solution of problem $(AVIP)$. Note that this result can be deduced from Theorem 2.4. Indeed, the asymmetric projection iteration is a particular case of subproblem $(AP^k)$ where $\Omega^k = \Omega = G - \lambda Q$, and $\lambda^k = \lambda$ for all $k$. Hence, we have to take

$$h(x) = (1/2)x^T Gx \text{ and } L = -Q$$

in the decomposition of $\Omega^k$ considered in Theorem 2.4. Then in this case, Condition $(RC)$ is satisfied since $F(x) - L(x) = (1/2)(M + M^T)x - q$, and $M + M^T$ has the Dunn property since it is positive semi-definite and symmetric (see Proposition 1.28).

Moreover, conditions ensuring the existence of a solution of problem $(AVIP)$ are proposed by Goeleven et al. They are based on the concept of recession cone. Let $C$ be a nonempty subset of $I\!R^n$, the recession cone of $C$ is the closed cone $C_\infty$ defined by: $x$ belongs to $C_\infty$ if and only if there exist sequences $\{t_n\}_{n\in I\!N} > 0$ and $\{\chi_n\}_{n\in I\!N} \subset C$ such that $t_n \to +\infty, t_n^{-1}\chi_n \to x$. If $C$ is closed and convex, then

$$C_\infty = \cap_{\lambda>0} \frac{C - x_0}{\lambda},$$

where $x_0$ is any element in $C$.

Let $M$ be a general positive semi-definite matrix. Then the existence of at least one solution of problem $(AVIP)$ is guaranteed (see [53]) provided that

$$\exists\, x_0 \in C, \forall x \in \mathrm{Ker}(M + M^T) \cap C_\infty \setminus \{0\} : \langle q - Mx_0, x\rangle < 0.$$

When $M$ has the Dunn property, this condition can be replaced by the less restrictive one (see [55], Theorem 2.2):

$$\langle q, x\rangle < 0,\ \forall x \in \mathrm{Ker}(M) \cap C_\infty \setminus \{0\}.$$

**Example 2.1** Let $C = \{(x_1, x_2) \in I\!R^2 : x_1 > 0, x_1x_2 \geq 1\}$, and let $M$ be defined by

$$M = \begin{pmatrix} a & -ka \\ -ka & k^2a \end{pmatrix},$$

with $a, k > 0$. The matrix $M$ is symmetric, positive semi-definite and has the Dunn property over $C$ with modulus $\gamma = ((k^2 + 1)a)^{-1} > 0$. Moreover, we have that $\mathrm{Ker}(M) = \{(x_1, x_2) : x_1 = kx_2\}$, and thus the condition "$\langle q, x\rangle < 0,\ \forall x \in \mathrm{Ker}(M) \cap I\!R^2_+ \setminus \{(0, 0)\}$" is equivalent to "$kq_1 + q_2 < 0$".

These solvability conditions are discussed in details in the applications considered in [55].

# Chapter 3

# Perturbation Methods

In the first section, we introduce and illustrate the convergence notion of Mosco for convex sets and functions and we discuss its use to approximate solutions of the original problem. In the second section, this variational convergence notion is coupled with the auxiliary problem principle to generate a general family of perturbation methods. We recall the most representative results existing for some particular instances of this framework and comment their weaknesses. In the last section, we introduce and motivate a rate of convergence notion for convex functions that will be used in subsequent chapters.

## 3.1 Convergence of Mosco and Approximation of Variational Inequalities

In the optimization setting and more generally to solve variational inequality problems, one often uses an algorithmically generated sequence of approximations of the admissible set, objective function or operator to obtain approximated subproblems with better computational properties. The approximations have to be constructed such that the sequence of points generated by the procedure accumulates at a solution of the original problem. We call variational convergence a concept which is particularly well designed for that purpose. A well-known notion of variational convergence is due to Mosco (see Ref. [87]) and originates from difficulties encountered in the approximation of the solutions of variational inequalities. First, we recall the Mosco–convergence for closed convex subsets of $H$.

**Definition 3.1** (See Ref. [5], Definition–Proposition 3.21) *Let $C^k, C$ be nonempty closed convex subsets of $H$. The sequence $\{C^k\}_{k\in\mathbb{N}}$ is said to be Mosco–convergent to $C$, which is denoted shortly by $C^k \xrightarrow{M} C$, if and only if the two following conditions hold:*

**(i)** *if for some sequence $\{k_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$, $w^{k_n} \in C^{k_n}$ for all $k_n$ and $\{w^{k_n}\}_{n\in\mathbb{N}}$ weakly converges to $w$, then one has $w \in C$;*

**(ii)** *for all $w \in C$, there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ such that $w^k \in C^k$ for all $k$ and $\{w^k\}_{k\in\mathbb{N}}$ strongly converges to $w$.*

The geometric interpretation of the Mosco–convergence of convex sets is given in the following proposition.

**Proposition 3.1** (See [5], Theorem 3.33) *Let $\{C^k\}_{k\in\mathbb{N}}, C$ be nonempty closed convex subsets of $H$. The following statements are equivalent:*

**(i)** $C^k \xrightarrow{M} C$,

**(ii)** $\forall x \in H: \ \mathit{proj}_{C^k}\, x \to \ \mathit{proj}_C\, x$,

**(iii)** $\forall x \in H: \ \mathit{dist}(x, C^k) \to \ \mathit{dist}(x, C)$.

Mosco also defines a convergence notion for convex functions by saying that a sequence of functions $\{\varphi^k\}_{k\in\mathbb{N}}$ converges to $\varphi$ if the sequence of epigraphs of $\varphi^k$ converges in the sense of Definition 3.1 to the epigraph of $\varphi$. This notion can be characterized in the following way:

**Definition 3.2** (See [87], Definition 1.4, Lemma 1.10) *Let $\{\varphi^k\}_{k\in\mathbb{N}}, \varphi \in \Gamma_0(H)$. The sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ is Mosco–convergent to $\varphi$, which we denote by $\varphi^k \xrightarrow{M} \varphi$, if and only if the two following conditions hold for all $w \in H$:*

**(i)** *for every sequence $\{w^k\}_{k\in\mathbb{N}}$ weakly converging to $w$, one has*

$$\underline{\lim}_{k\to\infty} \varphi^k(w^k) \geq \varphi(w);$$

**(ii)** *there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ strongly converging to $w$ such that*

$$\overline{\lim}_{k\to\infty} \varphi^k(w^k) \leq \varphi(w).$$

**Remark 3.1** It follows from this definition that if $\varphi^k \xrightarrow{M} \varphi$, then any $w \in H$ is the limit in the strong topology of a sequence $\{w^k\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \varphi^k(w^k) = \varphi(w).$$

Observe also that when $\varphi \leq \varphi^k$ for all $k$, condition (i) obviously holds. Note that in finite dimension, the concept of Mosco–convergence of functions coincide with that of epiconvergence (see, for example, [7], [5]).

The Mosco–convergence of a sequence of functions is connected to the simple convergence of the Moreau–Yosida approximates by the following proposition.

**Proposition 3.2** (See [5], Theorem 3.26) *Let* $\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$.

$$\varphi^k \overset{M}{\to} \varphi \Leftrightarrow \forall\lambda > 0,\, \forall x \in H,\, \varphi^k_\lambda(x) \to \varphi_\lambda(x).$$

Recall that the Moreau–Yosida approximate of a function $\varphi \in \Gamma_0(H)$ is defined for all $\lambda > 0$ and all $x \in H$ by:

$$\varphi_\lambda(x) = \inf_{y\in H}\{\varphi(y) + (1/2\lambda)\|y - x\|^2\}.$$

Based on this property, some authors use the following semi–distance between two functions $\varphi^1, \varphi^2 \in \Gamma_0(H)$ introduced in [8]:

**Definition 3.3** *Let* $\varphi^1, \varphi^2 \in \Gamma_0(H)$*. For all* $\lambda > 0$ *and all* $\rho \geq 0$*, we define:*

$$d_{\lambda,\rho}(\varphi^1, \varphi^2) = \sup_{\|x\|\leq\rho} |\varphi^1_\lambda(x) - \varphi^2_\lambda(x)|.$$

The Mosco–convergence is then connected to the semi–distance $\mathrm{d}_{\lambda,\rho}$ $(\lambda > 0, \rho \geq 0)$ by the following proposition:

**Proposition 3.3** (See [8], Theorem 2.51) *Let* $\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$*. If there exists* $\lambda_0 > 0$ *such that*

$$\lim_{k\to\infty} d_{\lambda_0,\rho}(\varphi^k, \varphi) = 0,\, \forall\rho \geq 0,$$

*then* $\varphi^k \overset{M}{\to} \varphi$*. The converse is true when* $H$ *is finite dimensional.*

The use of the Mosco–convergence in the optimization field is based on the following property: if $\varphi^k \overset{M}{\to} \varphi$, then the sequence $\{u_k\}_{k\in I\!N}$ defined by $u^k \in \operatorname{argmin}_{x\in H} \varphi^k(x)$ for each $k$, weakly accumulates at a solution $\bar{u} \in \operatorname{argmin}_{x\in H} \varphi(x)$ (see [5], Theorem 1.10). This result will be generalized for problem $(GVIP)$. Adopting the point of view of variational convergence, problem $(GVIP)$ can be perturbed by replacing the original function $\varphi$ by an approximate function $\varphi^k$ to get the new problem:

$$(PGVIP^k)\begin{cases} \text{find } u^k \in H \text{ and } r(u^k) \in F(u^k) \text{ such that, for all } x \in H, \\ \langle r(u^k), x - u^k\rangle + \varphi^k(x) - \varphi^k(u^k) \geq 0. \end{cases}$$

Assumptions ensuring that the sequence $\{u_k\}_{k\in I\!N}$ weakly accumulates at a solution of problem $(GVIP)$ will be set in Theorem 3.1 below but to prove it we need the following lemma:

**Lemma 3.1** *Let $\varphi \in \Gamma_0(H)$ and let $F$ be an upper hemi-continuous multivalued operator with nonempty convex and weakly compact values. If for some $\bar{u} \in dom\ \varphi$,*

$$\langle r(y), y - \bar{u}\rangle + \varphi(y) - \varphi(\bar{u}) \geq 0,\ \forall y \in dom\ \varphi,\ \forall r(y) \in F(y),$$

*then there exists some $\bar{r} \in F(\bar{u})$ such that*

$$\langle \bar{r}, y - \bar{u}\rangle + \varphi(y) - \varphi(\bar{u}) \geq 0, \forall y \in dom\ \varphi,$$

*which means that $\bar{u}$ is a solution of problem $(GVIP)$.*

**Proof.** The proof will be composed of two parts.
• Firstly, we prove that condition (3.1) here below is sufficient to ensure that $\bar{u}$ is a solution of problem $(GVIP)$:

$$\forall y \in \text{dom } \varphi,\ \inf_{r\in F(\bar{u})} \{\, \langle r, \bar{u} - y\rangle + \varphi(\bar{u}) - \varphi(y)\,\} \leq 0. \tag{3.1}$$

So, let us suppose that $\bar{u}$ satisfies (3.1). Then, we have that

$$\sup_{y\in\text{dom }\varphi}\ \inf_{r\in F(\bar{u})} \{\, \langle r, \bar{u} - y\rangle + \varphi(\bar{u}) - \varphi(y)\,\} \leq 0.$$

Since dom $\varphi$ is nonempty and convex, $F(\bar{u})$ is nonempty, convex and weakly compact, the function $f(r, y) = \langle r, \bar{u} - y\rangle + \varphi(\bar{u}) - \varphi(y)$ is linear in $r$ and concave in $y$, we can apply the classical minsup theorem (see, for example, [117] p.433 or also [9], Chapter 2, Section 7, Theorem 1) to obtain that

$$\begin{aligned} &\sup\nolimits_{y\in\text{dom }\varphi}\ \min\nolimits_{r\in F(\bar{u})} \{\, \langle r, \bar{u} - y\rangle + \varphi(\bar{u}) - \varphi(y)\,\} \\ &= \min\nolimits_{r\in F(\bar{u})}\ \sup\nolimits_{y\in\text{dom }\varphi} \{\, \langle r, \bar{u} - y\rangle + \varphi(\bar{u}) - \varphi(y)\,\}. \end{aligned}$$

And hence, there exists $\bar{r} \in F(\bar{u})$ such that

$$\sup_{y \in \mathrm{dom}\ \varphi} \{ \langle \bar{r}, \bar{u} - y \rangle + \varphi(\bar{u}) - \varphi(y) \} \leq 0.$$

Therefore,

$$\langle \bar{r}, \bar{u} - y \rangle + \varphi(\bar{u}) - \varphi(y) \leq 0,\ \forall y \in \mathrm{dom}\ \varphi$$

i.e., $\bar{u}$ solves problem $(GVIP)$.

• Secondly, we prove the lemma by contradiction. For that, let us assume that $\bar{u}$ does not solve problem $(GVIP)$. By the first part of the proof, it follows that $\bar{u}$ does not satisfy (3.1) so that there exists an element $\bar{y}$ in dom $\varphi$ such that

$$\langle r, \bar{u} - \bar{y} \rangle + \varphi(\bar{u}) - \varphi(\bar{y}) > 0,\ \forall r \in F(\bar{u}). \tag{3.2}$$

For all $\lambda \in ]0,1]$, set $\bar{y}_\lambda = \bar{u} + \lambda(\bar{y} - \bar{u})$ and consider the mapping $\gamma$:

$$\lambda \to \gamma(\lambda) = \{ \langle r_\lambda, \bar{u} - \bar{y} \rangle + \varphi(\bar{u}) - \varphi(\bar{y}) : r_\lambda \in F(\bar{y}_\lambda) \}.$$

From (3.2), $I\!R_0^+$ is a neighborhood of the set $\gamma(0)$. Moreover, since $F$ is upper hemi-continuous at $\bar{u}$, we have that $\gamma$ is upper semi-continuous at $0^+$, and thus that $\gamma(\tilde{\lambda}) \subset I\!R_0^+$ for all $\tilde{\lambda} \in ]0,1]$ sufficiently small. This means that

$$\langle r_{\tilde{\lambda}}, \bar{u} - \bar{y} \rangle + \varphi(\bar{u}) - \varphi(\bar{y}) > 0,\ \forall\, r_{\tilde{\lambda}} \in F(\bar{y}_{\tilde{\lambda}}).$$

From the convexity of $\varphi$, we derive that

$$\langle r_{\tilde{\lambda}}, \bar{u} - \bar{y}_{\tilde{\lambda}} \rangle + \varphi(\bar{u}) - \varphi(\bar{y}_{\tilde{\lambda}}) > 0,\ \forall\, r_{\tilde{\lambda}} \in F(\bar{y}_{\tilde{\lambda}}).$$

And this contradicts the hypothesis. □

In the case where $\varphi$ is the indicator function of a nonempty closed convex subset of $H$, this lemma reduces to Proposition 1 of [36], or Proposition 2.2 of [70].

Recall that an operator $F$ is locally hemibounded at a point $x \in$ dom $F$ if for each $y \in$ dom $F$, $y \neq x$, the element $x + \lambda(y - x) \in$ dom $F$ for any $\lambda \in [0,1]$ and the set $\cup_{0<\lambda\leq\bar{\lambda}} F(x + \lambda(y - x))$ is bounded in $H$. Hence, if $F$ is maximal monotone and locally hemibounded on dom $F$, it follows from Proposition 1.14 that $F$ is upper hemi-continuous.

**Theorem 3.1** *Assume that problems* $(GVIP)$ *and* $(PGVIP^k)$ *admit at least one solution and that the sequence* $\{u^k\}_{k\in\mathbb{N}}$ *generated by solving subproblems* $(PGVIP^k)$ *is bounded. If* $\{\varphi^k\}_{k\in\mathbb{N}}, \varphi \in \Gamma_0(H)$, $\varphi^k \xrightarrow{M} \varphi$ *and one of the following conditions is satisfied:*

**(1)** $F$ *is such that if a sequence* $\{\chi^k\}_{k\in\mathbb{N}}$ *weakly converges to some* $\bar{\chi}$ *and* $r^k \in F(\chi^k)$ *for all* $k$*, then the sequence* $\{r^k\}_{k\in\mathbb{N}}$ *strongly converges to some* $\bar{r} \in F(\bar{\chi})$,

**(2)** $F$ *is monotone, bounded, upper hemi-continuous and has nonempty convex weakly compact values,*

*then each weak limit point of the sequence* $\{u^k\}_{k\in\mathbb{N}}$ *is a solution of problem* $(GVIP)$.

**Proof.** Let $\bar{u}$ be a weak limit point of the sequence $\{u^k\}_{k\in\mathbb{N}}$ and let $\{u^k\}_{k\in K\subset\mathbb{N}}$ be a subsequence weakly converging to $\bar{u}$. Since $\varphi^k \xrightarrow{M} \varphi$, we have that

$$\underline{\lim}_{k\in K}\varphi^k(u^k) \geq \varphi(\bar{u}).$$

Moreover, by Remark 3.1, for each $y \in \text{dom } \varphi$, there exists a sequence $\{y^k\}_{k\in\mathbb{N}}$ such that

$$y^k \to y \text{ and } \varphi^k(y^k) \to \varphi(y).$$

By definition of $\{u^k\}_{k\in\mathbb{N}}$, we have that for each $k$, there exists $r(u^k) \in F(u^k)$ such that

$$\langle r(u^k), y^k - u^k\rangle + \varphi^k(y^k) - \varphi^k(u^k) \geq 0. \tag{3.3}$$

Hence, if $F$ satisfies Conditon (1) and if we pass to the superior limit on $k \in K$ in inequality (3.3), then we obtain that there exists some $\bar{r} \in F(\bar{u})$ such that

$$\langle \bar{r}, y - \bar{u}\rangle + \varphi(y) - \varphi(\bar{u}) \geq 0,$$

which means that $\bar{u}$ is a solution of problem $(GVIP)$.

Now, when Condition (2) holds, it follows from the monotonicity of $F$ that, for any $r(y) \in F(y)$,

$$\begin{aligned}\langle r(u^k), y^k - u^k\rangle &= \langle r(u^k), y^k - y\rangle + \langle r(u^k) - r(y), y - u^k\rangle + \langle r(y), y - u^k\rangle \\ &\leq \langle r(u^k), y^k - y\rangle + \langle r(y), y - u^k\rangle.\end{aligned}$$

So that, from the boundedness of $F$, the sequence $\{r(u^k)\}_{k\in I\!N}$ is bounded and we get

$$\overline{\lim}_{k\in K}\langle r(u^k), y^k - u^k\rangle \leq \langle r(y), y - \bar{u}\rangle.$$

Now, passing to the superior limit on $k \in K$ in (3.3), we deduce that

$$\langle r(y), y - \bar{u}\rangle + \varphi(y) - \varphi(\bar{u}) \geq 0.$$

We can then apply lemma 3.1 and conclude that $\bar{u}$ is a solution of problem $(GVIP)$. $\square$

Remark that in this theorem, the existence of a solution of the approximate problems $(PGVIP^k)$ and boundedness of the sequence of approximating solutions have to be checked. When $F$ is singlevalued, coerciveness of the operator $F$ guarantees the existence of a solution of the original problem $(GVIP)$ as well as the approximate problems $(PGVIP^k)$. For example, Mosco proves in [87] (Theorem B) that if F is a strongly monotone, bounded and hemi-continuous operator, then the sequence generated by solving subproblems $(PGVIP^k)$ converges to the unique solution of $(GVIP)$. The situation is more critical when the operator $F$ is not coercive and when the solution of problem $(GVIP)$ and of the subproblems $(PGVIP^k)$ can be multiple. To remedy this problem, Mosco makes use of the so–called "elliptic regularization". It consists in adding to the operator $F$ the coercive perturbation $k^{-\alpha}I(\alpha > 0)$ and in considering the regularized approximated problems:

$$(RPGVIP^k)\begin{cases} \text{find } \tilde{u}^k \in H \text{ such that, for all } x \in H, \\ \langle F(\tilde{u}^k) + k^{-\alpha}\tilde{u}^k, x - \tilde{u}^k\rangle + \varphi^k(x) - \varphi^k(\tilde{u}^k) \geq 0. \end{cases}$$

The approach which is then used by Mosco requires an approximation of the original function $\varphi$ by the sequence $\{\varphi^k\}_{k\in I\!N}$ in a stronger sense defined by the following properties:

**Condition (Mo):**
*there exists a positive constant $\alpha$ such that*

**(i)** *for any weakly convergent sequence $\{w^k\}_{k\in I\!N}$ with $\overline{\lim}_{k\to\infty}\varphi^k(w^k) < +\infty$, there exists a subsequence $\{\varphi^{k_n}(w^{k_n})\}_{k_n\in K\in I\!N} \subset \{\varphi^k(w^k)\}_{k\in I\!N}$ and a sequence $\{v^n\}_{n\in I\!N}$ such that*

$$k_n^\alpha(w^{k_n} - v^n) \rightharpoonup 0 \text{ and } \underline{\lim}_{k\to\infty} k_n^\alpha(\varphi^{k_n}(w^{k_n}) - \varphi(v^n)) \geq 0;$$

**(ii)** *for any $w \in H$, there exists a sequence $\{w^k\}_{k\in I\!N}$ such that*

$$k^\alpha(w^k - w) \to 0 \text{ and } \overline{\lim}_{k\to\infty} k^\alpha(\varphi^k(w^k) - \varphi(w)) \leq 0.$$

When $\varphi \leq \varphi^k$ for all $k$, condition (i) holds. To see this, it suffices to choose $v^n = w^{k_n}$ for each $n$.

When $\varphi = \Psi_C, \varphi^k = \Psi_{C^k}$, with $C, C^k$ nonempty closed convex subsets of $H$, Goeleven and Salmon (see [54]) propose a similar approach. They also regularize the mapping $F$ by $F + \epsilon_k I$ where $\{\epsilon_k\}_{k\in I\!N}$ is a sequence of positive numbers converging to zero, but they replace the strong convergence Condition (Mo) on the sequence $\{\varphi^k\}_{k\in I\!N}$ by a relabelling of the sequence $\{C^k\}_{k\in I\!N}$ to obtain the following result:

**Proposition 3.4** (See [54], Theorem 2.2.1) *Assume that problem $(VIP)$ admits at least one solution. If $\{C^k\}_{k\in I\!N}, C$ are nonempty closed convex subsets of $H$, $C^k \xrightarrow{M} C$, $\epsilon^k \to 0^+$, $F$ is a maximal monotone and Lipschitz continuous operator, then there exists an increasing function $i : I\!N \to I\!N$ such that the sequence $\{\bar{u}^k\}_{k\in I\!N}$ generated by solving the problem:*

$$\begin{cases} \text{find } \bar{u}^k \in C^{i(k)} \text{ such that, for all } x \in C^{i(k)}, \\ \langle F(\bar{u}^k) + \epsilon_k \bar{u}^k, x - \bar{u}^k\rangle \geq 0, \end{cases}$$

*strongly converges to the solution of miniminal norm of problem $(VIP)$.*

Moreover, in their convergence analysis of discretization methods for complementarity problems, Goeleven and Salmon ([54]) specialize the concept of Mosco–convergence by what they call the Glowinski–convergence (see [54], Definition 1.3). Exploiting this concept of convergence and using recent tools of recession analysis, they prove weak convergence of the sequence generated by solving the regularized subproblems $(RPGVIP^k)$ to the solution of minimal norm of the complementarity problem when the operator $F$ is supposed to be monotone, bounded and hemi-continuous (see [54], Theorem 2.1.1).

Let us now illustrate the concept of Mosco–convergence by some examples:

**Example 3.1 Monotone Sequence of Functions**

**Proposition 3.5** (See [5], Theorem 3.20) *Let* $\{\varphi^k\}_{k\in I\!N} \in \Gamma_0(H)$.

**(i)** *If the sequence* $\{\varphi^k\}_{k\in I\!N}$ *is monotonically increasing, then*

$$\varphi^k \xrightarrow{M} \sup_{k\in I\!N} \varphi^k.$$

**(ii)** *If the sequence* $\{\varphi^k\}_{k\in I\!N}$ *is monotonically decreasing, then*

$$\varphi^k \xrightarrow{M} cl(\inf_{k\in I\!N} \varphi^k).$$

More particularly, when $\varphi$ is the indicator function of a nonempty closed convex subset $C$ of $H$, we can deduce from Proposition 3.5 the following result concerning the approximation of $C$ by a sequence $\{C^k\}_{k\in I\!N}$ of subsets of $H$:

**Proposition 3.6** (See [5], Theorem 3.22) *Let* $\{C^k\}_{k\in I\!N}, C$ *be nonempty closed convex subsets of* $H$.

**(i)** *(exterior approximation of* $C$*)*
*If* $C^{k+1} \subset C^k$ *for every* $k \in I\!N$, *and* $C = \cap_{k\in I\!N} C^k$, *then*

$$\Psi_{C^k} \xrightarrow{M} \Psi_C \ (i.e.\ C^k \xrightarrow{M} C).$$

**(ii)** *(interior approximation of* $C$*)*
*If* $C^k \subset C^{k+1}$ *for every* $k \in I\!N$, *and* $C = cl(\bigcup_{k\in I\!N} C^k)$, *then*

$$\Psi_{C^k} \xrightarrow{M} \Psi_C \ (i.e.\ C^k \xrightarrow{M} C).$$

**Example 3.2 Exterior Penalty Functions**

Let $g_i : H \to I\!R, i = 0, 1, ..., m$ be convex functions and let $C = \{x \in H : g_i(x) \leq 0, i = 1, ..., m\}$. We suppose that $C$ and the closure of its interior coincide. We take $\varphi = g_0 + \Psi_C$ and we consider the sequence of approximated functions $\varphi^k = g_0 + q^k$, where the sequence $\{q^k\}_{k\in I\!N}$ of lower semi-continuous convex real–valued functions belongs to the class of exterior penalty functions in the sense that it satisfies the following properties:

**(1)** $0 \leq q^k(x) \leq q^{k+1}(x)$, $\forall x \in H$, $\forall k \in I\!N$;

**(2)** if $x \in C$, then $q^k(x) = 0$, $\forall k \in I\!N$;

**(3)** if $x \notin C$, then $\lim_{k\to\infty} q^k(x) = +\infty$.

Well-known examples are the classical exterior penalty functions defined by

$$q^k(x) = (\nu_k/2)\sum_{i=1}^{m}[\max(0, g_i(x))]^2, \ \forall x \in H, \ \forall k \in I\!N,$$

and the exact exterior penalty functions defined by

$$q^k(x) = \nu_k \sum_{i=1}^{m} \max(0, g_i(x)), \ \forall x \in H, \ \forall k \in I\!N,$$

where $\{\nu_k\}_{k\in I\!N}$ is an increasing sequence of positive numbers converging to $+\infty$.

**Proposition 3.7** (See [125], Proposition VI.3.2) *If the sequence of penalty functions $\{q^k\}_{k\in I\!N}$ satisfies properties (1),(2),(3) above, then $\varphi^k \stackrel{M}{\rightarrow} \varphi$.*

Note that this result follows directly from Proposition 3.5 (i).

These penalties have been proven useful in the study of various optimization algorithms (see, for example, [13], [49], [73]). The drawback of such methods is that they generate points that are not in general in the feasible set $C$.

**Example 3.3 Barrier Functions**

Let $\varphi$ be defined like in the preceding example and consider the sequence of approximated functions $\varphi^k = g_0 + b(\nu_k, \cdot)$, where the sequence $\{\nu_k\}_{k\in I\!N}$ of positive barrier parameters is strictly increasing to $+\infty$ and the barrier function $b$ associated with $C$ is such that $b(\nu_k, \cdot)$ is continuous and positive on the interior of $C$, takes the value $+\infty$ elsewhere and, for each $x$ in the interior of $C$, the sequence $\{b(\nu_k, x)\}_{k\in I\!N}$ is strictly decreasing to 0. The most commonly used barrier functions associated with $C$ are:

$$b_1(\nu_k, x) = -\nu_k^{-1}\sum_{i=1}^{m}[\min(0, g_i(x))]^{-1}, \quad \forall x \in H, \quad \forall k \in I\!N; \tag{3.4}$$

$$b_2(\nu_k, x) = \nu_k^{-2}\sum_{i=1}^{m}[\min(0, g_i(x))]^{-2}, \quad \forall x \in H, \quad \forall k \in I\!N; \tag{3.5}$$

$$b_3(\nu_k, x) = -\nu_k^{-1}\sum_{i=1}^{m}\ln[\min(1/2, -g_i(x))], \quad \forall x \in H, \quad \forall k \in I\!N, \tag{3.6}$$

where, by convention, $\ln(a) = -\infty$ if $a \leq 0$.
For these examples, we have that $\varphi^k \overset{M}{\to} \varphi$ (see [7], Theorem 4).

Barrier functions are extensively used in the literature (see [50] for a survey).

**Example 3.4 Tykhonov Regularization**

Any function $\varphi \in \Gamma_0(H)$ can be approximated by the following Tykhonov regularization:

$$\varphi^k = \varphi + (t_k/2)\,||\cdot||^2, \quad \forall k \in I\!N,$$

where the sequence of positive numbers $\{t_k\}_{k \in I\!N}$ is stricly decreasing to 0. In this case, it can be easily verified that $\varphi^k \overset{M}{\to} \varphi$. This perturbation is used, for example, in [74], [91], [125].

**Example 3.5 Global Regularization of a Finite Maximum Function**

Let the function $\varphi$ be the finite maximum function

$$\varphi(x) = \max_{j=1...m} f_j(x), \quad \forall x \in I\!R^n,$$

where for each $j$, $f_j$ is a convex continuously differentiable function defined on $I\!R^n$. Note that the function $\varphi$ is generally nondifferentiable at points where the maximum is attained for more than one function $f_j$.

Denote by $U$ the unit–simplex in $I\!R^n$:

$$U = \{u \in I\!R^m : \sum_{j=1}^{m} u_j = 1, u_j \geq 0, \forall j = 1, ..., m\}.$$

Let $\Delta_k$ be a sequence of positive numbers converging to zero and $\{v^k\}_{k \in I\!N}$ be a sequence of the set $U$. For each $k$, we consider the regularized functions $\varphi^k$ introduced in [15],

$$\varphi^k(x) = \max_{u \in U}\{\langle u, f(x)\rangle - (\Delta_k/2)\|u - v^k\|^2\},$$

where $f(x) = (f_1(x), ..., f_m(x))^T$. It is shown in [15] that for all $\Delta_k > 0$ and for all $v^k \in U$, the function $\varphi_k$ is differentiable on $I\!R^n$ and satisfies the following inequality:

$$\varphi(x) - \Delta_k \leq \varphi^k(x) \leq \varphi(x), \quad \forall k \in I\!N, \quad \forall x \in I\!R^n. \tag{3.7}$$

It is then clear that $\varphi^k \stackrel{M}{\to} \varphi$.

Note that this global regularization is used in [15] to minimize the function $\varphi$. The method presented there generates by this way differentiable subproblems that are approximately solved by using a quasi–Newton method. A particular choice for the sequence $\{v^k\}_{k\in I\!N}$ is proposed and is proven to be optimal. More precisely, $\{v^k\}_{k\in I\!N}$ is iteratively constructed in the following way:

$$\begin{cases} v^0 \in U, \\ \\ v^{k+1} = u(v^k, x^k), \end{cases}$$

where $x^k$ is the current iterate at iteration $k$ and $u(v^k, x^k)$ is the element in $U$ verifying

$$\varphi^k(x^k) = \langle u(v^k, x^k), f(x^k)\rangle - (\Delta_k/2)\|u(v^k, x^k) - v^k\|^2.$$

Observe that in this case, the sequence $\{\varphi^k\}_{k\in I\!N}$ is not chosen a priori but each $\varphi^{k+1}$ depends on the current iterate $x^k$. However, this dependence is not involved explicitely to show the Mosco–convergence property thanks to the inequalities (3.7).

For more details about the theory of Mosco–convergence, examples and applications, we refer to [5], [7], [87], [120], [125].

Besides its use to approximate solutions of the original problem, variational convergence has also been proven useful in combination with a standard iterative method. The idea to couple variational convergence with iterative methods to solve convex optimization problems or monotone variational inequalities has already been extensively used to perturb proximal point algorithms (see, for example, [2], [13], [45], [64], [88], [89], [90], [91], [125]), splitting methods as the forward–backward scheme, the Peaceman–Rachford scheme, the Douglas–Rachford scheme, the $\theta-$scheme (see [45], [57]) and Tykhonov algorithms (see, for example, [88], [125]). For the auxiliary problem principle, the influence of a variational perturbation of $\varphi$ is studied by Lemaire ([73]) in the optimization field and by Sonntag ([120]) and Makler et al. ([82]) for variational inequalities. In the following section, we recall and comment the results obtained for the perturbed auxiliary problem method.

## 3.2 The Perturbed Auxiliary Problem

A general family of perturbation methods to solve problem $(GVIP)$ can be obtained by combining the variational convergence theory with the auxiliary problem principle. More precisely, at iteration $k$, the auxiliary subproblem $(AP^k)$ can be perturbed by replacing the original function $\varphi$ by an approximate function $\varphi^k$ to obtain the new problem

$$(PAP^k) \begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)), x - x^{k+1}\rangle \\ +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

If $\Omega^k$ is symmetric ($\Omega^k = \nabla K^k$ for all $k$), the perturbed problem associated with $(SAP^k)$ will be denoted by $(PSAP^k)$.

The convergence study of this perturbed scheme has already been initiated in the literature for some particular instances. In the following, we recall the results that have been obtained before our own contribution.

In the case where $F$ is singlevalued, Makler et al. ([82]) have obtained the following convergence result for the scheme defined by subproblems $(PSAP^k)$ i.e., with the auxiliary operators $\Omega^k$ being the gradient of some continuously differentiable function $K^k$.

**Theorem 3.2** (See [82], Theorem 4.1) *Assume that the following conditions are satisfied:*

- $\{K^k\}_{k\in \mathbb{N}}$ *is a sequence of functions from $H$ into $\mathbb{R}$ which are continuously differentiable and strongly convex with modulus $\beta_k \geq \beta > 0$ over dom $\varphi$ for all $k \in \mathbb{N}$;*
- $\{\nabla K^k\}_{k\in \mathbb{N}}$ *is a sequence of Lipschitz continuous mappings with Lipschitz constants $\Lambda_k \leq \Lambda$ over dom $\varphi$ for all $k \in \mathbb{N}$;*
- *for each $k \in \mathbb{N}$, there exists a positive number $\eta_k$ such that, for all $x, y \in$ dom $\varphi$:*

$$K^{k+1}(x) - K^{k+1}(y) - \langle \nabla K^{k+1}(y), x - y\rangle$$
$$\leq \eta_k(K^k(x) - K^k(y) - \langle \nabla K^k(y), x - y\rangle);$$

- *there exist two positive numbers $\beta'$ and $\Lambda'$ such that, for all $k \in I\!N$,*

$$\beta_k/M^k \geq \beta' \text{ and } \Lambda_k/M^k \leq \Lambda',$$

  *where $M^k = \prod_{j=0}^{k-1} \eta_j$;*

- *$F$ is strongly monotone with modulus $\alpha > 0$ and Lipschitz continuous with Lipschitz constant $L > 0$;*

- *there exist $\underline{\lambda}$ and $\overline{\lambda}$ such that, for all $k \in I\!N$,*

$$0 < \underline{\lambda} \leq \lambda_k/M^k \leq \overline{\lambda} < 2\alpha\beta'/L^2;$$

- *$\{\varphi^k\}_{k \in I\!N}, \varphi \in \Gamma_0(H)$ are such that $\varphi^k \xrightarrow{M} \varphi$ and $\varphi \leq \varphi^k$ for all $k$.*

*Then the original problem $(GVIP)$ admits a unique solution $x^*$ and the sequence $\{x^k\}_{k \in I\!N}$ generated by solving subproblems $(PSAP^k)$ strongly converges to $x^*$.*

Note that the condition $\varphi \leq \varphi^k$ for all $k \in I\!N$, implies that the effective domains of the approximate functions $\varphi^k$ are contained in the domain of $\varphi$ so that interior approximations of the function $\varphi$ can be considered.

In the particular situation where $\varphi$ is the indicator function of a closed convex set $C$, $\varphi^k$ is the indicator function of a closed convex approximating subset of $C$ and $K^k(x) = (1/2)\|x\|^2$ for all $x \in H$ and all $k \in I\!N$, Theorem 3.2 reduces to the convergence result proposed by Sonntag in [120].

In the optimization field, Lemaire (see [73]) relaxed the strong convexity assumption by taking advantage of the particular structure of the problem. For solving the optimization problem $(OP)$ with $f$ a finite–valued differentiable function, he considers the following iteration scheme:

$$x^{k+1} = (I + \lambda_k \partial\varphi^k)^{-1}(x^k - \lambda_k \nabla f(x^k)).$$

It is clear that this problem is a particular instance of subproblem $(PAP^k)$ where $\Omega^k$ is the identity mapping for all $k$. The convergence result is then the following:

**Theorem 3.3** (See [73], Theorem 2) *Assume that problem $(OP)$ admits at least one solution and that the following conditions are satisfied:*

- *$f : H \to I\!R$ is convex and differentiable;*
- *$\nabla f$ is Lipschitz continuous on dom $\varphi$ with Lipschitz constant $L$;*
- *there exists $\underline{\lambda}, \overline{\lambda} > 0$ such that, for all $k \in I\!N$,*

$$0 < \underline{\lambda} \leq \lambda_k \leq \overline{\lambda} < 2/L;$$

- *$\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ are such that $\varphi^{k+1} \leq \varphi^k$ for all $k$ and $\varphi = cl(\inf_{k\in I\!N} \varphi^k)$;*
- *$\{x^k\}_{k\in I\!N}$ is bounded.*

*Then provided that $x^0 \in$ dom $\varphi$, any weak limit point of the sequence $\{x^k\}_{k\in I\!N}$ is a solution of problem $(OP)$ and $\lim_{k\to+\infty}(f+\varphi^k)(x^k) = \inf(f+\varphi)$.*

The assumption that the sequence $\{x^k\}_{k\in I\!N}$ is bounded is crucial. Nevertheless, Lemaire proves that it is satisfied when the domain of $\varphi$ is bounded or $f+\varphi$ is coercive (see [73], Proposition 4). Note that the condition on the sequence $\{\varphi^k\}_{k\in I\!N}$ implies that $\varphi^k \xrightarrow{M} \varphi$ (see Proposition 3.5(ii)). Moreover, Lemaire shows that this condition can be replaced by the following in terms of the variational semi–distance between $\varphi^k$ and $\varphi$:

$$\forall \rho \geq 0,\ \mathrm{d}_{\underline{\lambda},\rho}(\varphi^k, \varphi) \to 0.$$

This condition implies also that $\varphi^k \xrightarrow{M} \varphi$ and is equivalent to the Mosco–convergence if $H$ is finite dimensional (see Proposition 3.3).

The weaknesses and lacknesses of Theorems 3.2 and 3.3 have already been discussed in the introduction. For the general setting of variational inequalities, Theorem 3.2 suffers from the facts that it is restricted to a strongly monotone, Lipschitz continuous, singlevalued mapping and to symmetric auxiliary operators. In Chapter 4, we present convergence results for the perturbed auxiliary scheme with nonsymmetric auxiliary operators when $F$ is singlevalued and satisfies the same kind of assumptions as in the case where $\varphi$ is not perturbed. In Chapter 5, we study the multivalued case i.e., we extend the results of Section 2.2.2 to allow some perturbations of the function $\varphi$.

## 3.3 $\alpha$–Order Convergence

To obtain convergence results for the perturbed auxiliary problem scheme under weaker assumptions on the mapping $F$ than strong monotonicity, we need to control the rapidity of convergence of the approximations $\varphi^k$ to $\varphi$. So, adding a condition on the speed of convergence of the sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ is the price to pay to achieve our goal. The notion of "$\alpha$–order" convergence is defined here below.

**Definition 3.4** *Let $\{\varphi^k\}_{k\in\mathbb{N}}, \varphi \in \Gamma_0(H), w \in H$, and $\alpha > 1$. The sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ is said to converge to $\varphi$ at the order $\alpha$ at $w$ if there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ such that*

$$k^\alpha \|w^k - w\| \to 0 \text{ and } k^\alpha |\varphi^k(w^k) - \varphi(w)| \to 0. \tag{3.8}$$

A similar notion is used by Mosco in [87] to deal with the case where the operator in the variational inequality is noncoercive (see Condition $(Mo)$).

So, in addition to the Mosco–convergence of the sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ to $\varphi$, we will assume that, for some solution $x^*$ to the original problem, there exists a constant $\alpha > 1$ such that the sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ converges to $\varphi$ at the order $\alpha$ at $x^*$. This will be used in the proofs mainly to ensure that the series

$$\sum_{k=0}^{+\infty} \|w^k - x^*\| \text{ and } \sum_{k=0}^{+\infty} |\varphi^k(w^k) - \varphi(x^*)|$$

are convergent.

Let us now see on some examples what this speed of convergence amounts to.

**Example 3.6 Interior Approximation of the Feasible Set**

Consider that we want to approximate $\varphi = \Psi_C$, where $C$ is a nonempty closed convex subset of $H$, by the sequence of indicator functions $\{\varphi^k\}_{k\in\mathbb{N}} = \{\Psi_{C^k}\}_{k\in\mathbb{N}}$, where $\{C^k\}_{k\in\mathbb{N}}$ is a sequence of nonempty closed convex subsets of $C$. As already observed in Proposition 3.6(ii), $\Psi_{C^k} \xrightarrow{M} \Psi_C$ when the sequence $\{C^k\}_{k\in\mathbb{N}}$ is chosen in such a way that $C = \text{cl}(\bigcup C^k)$ and $C^k \subset C^{k+1}$, for all $k \in \mathbb{N}$. If, in addition, we impose that there exist $\alpha > 1$ and $\beta > 0$ such that, for all $k \in \mathbb{N}$,

$$\max_{y\in C} \min_{x\in C^k} \|x - y\| \leq 1/k^{\alpha+\beta},$$

then the sequence $\{\Psi_{C^k}\}_{k\in I\!N}$ converges to $\Psi_C$ at the order $\alpha$ at any $w \in C$. To see this, it suffices to take $w^k = \mathrm{proj}_{C^k} w$ in (3.8).

**Example 3.7 Barrier Functions**

Let $\varphi^k, \varphi$ be as in Example 3.3. The next proposition shows that the $\alpha$–order convergence is reached at any $w \in C$ provided that the barrier parameters increase sufficiently fast to $+\infty$.

**Proposition 3.8** *If there exist $\alpha > 1$ and $\beta > 0$ such that $k^{\alpha+\beta}/\nu_k \to 0$, then the functions $\varphi^k = g_0 + b_3(\nu_k, \cdot)$ converge to $\varphi$ at the order $\alpha$ at any $w \in C$.*

**Proof.** Let $w \in C$ and let $\alpha > 1$ and $\beta > 0$ such that $k^{\alpha+\beta}/\nu_k \to 0$. Choose $\widetilde{w} \in$ int $C$ and for each $k \in I\!N$, let us consider the point

$$w^k = k^{-\alpha-\beta}\widetilde{w} + (1 - k^{-\alpha-\beta})w,$$

which belongs to the interior of $C$.
It is clear that $k^\alpha(w^k - w) \to 0$. So, relation (3.8) will be satisfied if we prove that $k^\alpha b_3(\nu_k, w^k) \to 0$ or equivalently, that for each $i = 1, \ldots, m$,

$$k^\alpha \nu_k^{-1} \ln[\min(1/2, -g_i(w^k))] \to 0. \tag{3.9}$$

Since $g_i(w^k) \to g_i(w)$, it is obvious that (3.9) is true when $g_i(w) < 0$. Let us suppose now that $g_i(w) = 0$. By convexity of $g_i$, we have, for all $k$ sufficiently large, that $0 < -g_i(w^k) \le 1/2$ and $g_i(w^k) \le k^{-\alpha-\beta} g_i(\widetilde{w})$.
Therefore we deduce immediately that for all $k$ large enough,

$$\ln(-k^{-\alpha-\beta} g_i(\widetilde{w})) \le \ln[\min(1/2, -g_i(w^k))] \le -\ln 2.$$

Finally, multiplying each term of the above inequality by $k^\alpha \nu_k^{-1}$ and passing to the limit on $k$, we obtain easily (3.9) provided we observe that $k^\alpha/\nu_k$ and $k^\alpha \ln(k^{\alpha+\beta})/\nu_k$ tend to zero when $k \to +\infty$. $\square$

If we set $\nu_k = k^\gamma$ for all $k \in I\!N$, with $\gamma > 1$, then it is always possible to find $\alpha > 1$ and $\beta > 0$ such that $k^{\alpha+\beta}/\nu_k \to 0$.

**Example 3.8 Tykhonov Regularization**

For the Tykhonov regularization considered in Example 3.4, the $\alpha$–order convergence is satisfied at any $w \in H$ of the original problem provided that the parameter sequence $\{t_k\}_{k\in I\!N}$ converges fast enough to zero. Indeed,

$$k^\alpha|\varphi^k(w) - \varphi(w)| = (k^\alpha t_k/2)\|w\|^2.$$

So, if there exists $\alpha > 1$ such that $k^\alpha t_k \to 0$, then we can take $w^k = w$ for all $k$ in Definition 3.4.

**Example 3.9 Global Regularization of a Finite Maximum Function**

For the example treated in Example 3.5, we have the inequalities (3.7) with $\{\Delta_k\}_{k\in I\!N}$ being a sequence that converges to zero such that for each $w$,

$$k^\alpha|\varphi^k(w) - \varphi(w)| \leq k^\alpha \Delta_k.$$

So, if there exists $\alpha > 1$ such that $k^\alpha \Delta_k \to 0$, then we obtain the $\alpha$–order convergence of $\varphi^k$ to $\varphi$ at any $w$.

# Chapter 4

# Convergence of the Perturbed Auxiliary Problem Method for Singlevalued Mappings

In this chapter, the operator $F$ is assumed to be singlevalued. In that case, the perturbed auxiliary subproblem at iteration $k$ can be written

$$(PAP^k)\begin{cases} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)), x - x^{k+1}\rangle \\ +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

We focus our attention on interior approximations of the function $\varphi$ i.e., we assume that $\varphi \leq \varphi^k$ for all $k \in I\!N$, and we suppose that the convergence of $\{\varphi^k\}_{k\in I\!N}$ to $\varphi$ is sufficiently fast. Global and local convergence results for this family of perturbation methods are obtained under a kind of pseudo Dunn property which links the mapping $F$ and the nonsymmetric components of the auxiliary operators. Then we deduce from this convergence analysis, generalizations of the results mentioned in Chapter 2 Section 2.2.1 when $\varphi$ is not perturbed. The results of this part of the work generalize in the infinite dimensional case those appeared in [112].

## 4.1 Global Convergence Results

In this section, we present a global convergence analysis for the general scheme defined by problems $(PAP^k)$, $k \in I\!N$. In the sequel, we allow the sequence of auxiliary operators $\{\Omega^k\}_{k \in I\!N}$ to be built step by step. Indeed, very often, as for example in the Newton method, $\Omega^k$ depends explicitely on the iterate $x^k$. To emphasize this property, we will write $\Omega(x^k, \cdot)$ instead of $\Omega^k$. This new notation suggests us to introduce a function $\Omega$ from $H \times H$ into $H$ which will be useful to express assumptions independently of the iterate $x^k$. With this convention, problem $(PAP^k)$ can be written as:

$$(PAP^k) \left\{ \begin{array}{l} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F(x^k) + \lambda_k^{-1}(\Omega(x^k, x^{k+1}) - \Omega(x^k, x^k)), x - x^{k+1} \rangle \\ \hfill + \varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{array} \right.$$

Note that if $\varphi \leq \varphi^k$, the unique solution $x^{k+1}$ of this subproblem belongs to dom $\varphi$.

The next proposition is a first step towards the convergence study of the general iterative scheme. It highlights the minimal assumptions under which each weak limit point of the sequence $\{x^k\}_{k \in I\!N}$, if it exists, is a solution of problem $(GVIP)$.

**Proposition 4.1** *Assume that the following conditions hold true:*

- $F : H \to H$ *is singlevalued and weakly continuous on dom* $\varphi$ *and the functional* $x \to \langle F(x), x \rangle$ *is weakly lower semi-continuous on dom* $\varphi$*;*
- $\Omega(x, \cdot) : H \to H$ *is strongly monotone over dom* $\varphi$ *for each* $x \in$ *dom* $\varphi$ *and Lipschitz continuous over dom* $\varphi$ *uniformly in* $x$*;*
- $\lambda_k \geq \underline{\lambda} > 0$*, for all* $k \in I\!N$*;*
- $\{\varphi^k\}_{k \in I\!N}, \varphi \in \Gamma_0(H)$ *are such that* $\varphi^k \xrightarrow{M} \varphi$ *and* $\varphi \leq \varphi^k$ *for all* $k \in I\!N$*;*
- *the sequence* $\{x^k\}_{k \in I\!N}$ *is bounded and is such that the sequence* $\{\|x^{k+1} - x^k\|\}_{k \in I\!N}$ *converges to zero.*

*Then, provided that* $x^0$ *be chosen in dom* $\varphi$*, every weak limit point of the sequence* $\{x^k\}_{k \in I\!N}$ *is a solution of problem* $(GVIP)$*.*

**Proof.** Let $x^*$ be a weak limit point of $\{x^k\}_{k\in I\!N}$ and let $\{x^k\}_{k\in K\subset I\!N}$ be a subsequence weakly converging to $x^*$. Since $\{\|x^{k+1}-x^k\|\}_{k\in I\!N} \to 0$, we have that $\{x^{k+1}\}_{k\in K} \rightharpoonup x^*$ and also, from the fact that $\varphi^k \stackrel{M}{\to} \varphi$, that the following inequality holds:

$$\varphi(x^*) \leq \underline{\lim}_{k\in K}\varphi^k(x^{k+1}). \tag{4.1}$$

Moreover, for each $y$ in $H$, there exists a sequence $\{y^k\}_{k\in I\!N}$ such that

$$y^k \to y \quad \text{and} \quad \varphi^k(y^k) \to \varphi(y). \tag{4.2}$$

Using the following identity:

$$\begin{aligned}\langle F(x^k), y^k - x^{k+1}\rangle &= \langle F(x^k), y^k - y\rangle + \langle F(x^k), y\rangle \\ &\quad + \langle F(x^k), x^k - x^{k+1}\rangle - \langle F(x^k), x^k\rangle,\end{aligned}$$

and the fact that $F$ is weakly continuous on dom $\varphi$ and the function $x \to \langle F(x), x\rangle$ is weakly lower semi–continuous on dom $\varphi$, we obtain that

$$\overline{\lim}_{k\in K}\langle F(x^k), y^k - x^{k+1}\rangle \leq \langle F(x^*), y - x^*\rangle. \tag{4.3}$$

Now, by definition of $\{x^k\}_{k\in I\!N}$, we have that, for all $k \in I\!N$,

$$\begin{aligned}0 &\leq \langle F(x^k) + \lambda_k^{-1}(\Omega(x^k, x^{k+1}) - \Omega(x^k, x^k)), y^k - x^{k+1}\rangle \\ &\quad + \varphi^k(y^k) - \varphi^k(x^{k+1}).\end{aligned}$$

Passing then to the superior limit on $k \in K$ in the above inequality and using the Lipschitz continuity of $\Omega(x^k, \cdot)$ together with relations (4.1), (4.2), (4.3) and the fact that $\lambda_k > \underline{\lambda} > 0$, we obtain the following inequality:

$$0 \leq \langle F(x^*), y - x^*\rangle + \varphi(y) - \varphi(x^*),$$

which means that $x^*$ is a solution of $(GVIP)$. □

Observe that when $\varphi \leq \varphi^k$ for all $k$, the condition "$\varphi^k \stackrel{M}{\to} \varphi$" can be replaced by the following one:

$$\forall w \in H,\ \exists\{w^k\}_{k\in I\!N} : w^k \to w \text{ and } \varphi^k(w^k) \to \varphi(w).$$

When $F$ is weakly continuous on dom $\varphi$ then the function $x \to \langle F(x), x\rangle$ is weakly lower semi–continuous on dom $\varphi$, for example, if $F$ is compact on dom $\varphi$ or if $F$ is monotone on dom $\varphi$.

We have now to study under which assumptions the sequence $\{x^k\}_{k\in I\!N}$ is bounded and the sequence $\{\|x^{k+1} - x^k\|\}_{k\in I\!N}$ converges to zero. In the following result, we use a particular decomposition of $\Omega(x^k, y)$ :

$$\Omega(x^k, y) = \nabla h^k(y) + \lambda_k \; L(x^k, y), \quad \forall y \in H, \quad \forall k \in I\!N, \tag{4.4}$$

where $L$ is a singlevalued map from $H \times H$ into $H$ and $\{h^k\}_{k\in I\!N}$ is a sequence of functions from $H$ into $I\!R$. We assume that for all $x \in$ dom $\varphi$, $L(x, \cdot)$ is a monotone mapping over dom $\varphi$ and that for all $k \in I\!N$, $h^k$ is a continuously differentiable and strongly convex function over dom $\varphi$. Note that if the sequence $\{\Omega(x^k, \cdot)\}_{k\in I\!N}$ is uniformly strongly monotone with modulus $\beta > 0$, then it is always possible to decompose $\Omega(x^k, y)$ as in (4.4) by choosing $h^k(y) = (\beta/2)y^T y$ and $L(x^k, y) = \lambda_k^{-1}(\Omega(x^k, y) - \beta y)$ for all $x \in H$ and all $k \in I\!N$. With this formulation of $\Omega(x^k, y)$, problem $(PAP^k)$ can be written as:

$$(PAP^k) \begin{cases} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F(x^k) + L(x^k, x^{k+1}) - L(x^k, x^k) + \lambda_k^{-1}(\nabla h^k(x^{k+1}) - \nabla h^k(x^k)), \\ \qquad\qquad x - x^{k+1}\rangle + \varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

The two following lemmas will be used to prove the next proposition.

**Lemma 4.1** *For any two reals $a, b$ and any positive number $\tau$, we have*

$$ab \leq (\tau/2)\, a^2 + (1/2\tau)\, b^2.$$

**Lemma 4.2** (See [104], Lemma 3) *Let $\{u^k\}_{k\in I\!N}$, $\{s_k\}_{k\in I\!N}$, $\{t_k\}_{k\in I\!N}$ be non negative sequences satisfying*

$$u^{k+1} \leq s_k\, u^k + t_k, \; \forall k \in I\!N,$$

*and such that* $\lim_{k\to+\infty} \prod_{j=0}^{k} s_j > 0$ *and* $\sum_{k=0}^{+\infty} t_k < +\infty$, *then* $\{u^k\}_{k\in I\!N}$ *is a Cauchy sequence.*

**Proposition 4.2** *Assume that the solution set of problem $(GVIP)$ is nonempty and that the following conditions are satisfied:*

**(i)** *$\{h^k\}_{k\in I\!N}$ is a sequence of functions from $H$ into $I\!R$ which are continuously differentiable and strongly convex with modulus $\beta_k \geq \beta > 0$ over dom $\varphi$ for all $k \in I\!N$;*

**(ii)** $\{\nabla h^k\}_{k \in I\!N}$ *is a sequence of Lipschitz continuous mappings with Lipschitz constants* $\Lambda_k \leq \Lambda$ *over dom* $\varphi$ *for all* $k \in I\!N$*;*

**(iii)** *for each* $k \in I\!N$*, there exists a positive number* $\eta_k$ *such that, for all* $x, y \in$ *dom* $\varphi$*:*

$$h^{k+1}(x) - h^{k+1}(y) - \langle \nabla h^{k+1}(y), x - y \rangle$$

$$\leq \eta_k (h^k(x) - h^k(y) - \langle \nabla h^k(y), x - y \rangle);$$

**(iv)** *there exist two positive numbers* $\beta'$ *and* $\Lambda'$ *such that, for all* $k \in I\!N$*,*

$$\beta_k / M^k \geq \beta' \text{ and } \Lambda_k / M^k \leq \Lambda',$$

*where* $M^k = \prod_{j=0}^{k-1} \eta_j$*;*

**(v)** *there exist* $\underline{\lambda}$ *and* $\overline{\lambda}$ *such that, for all* $k \in I\!N$*,*

$$0 < \underline{\lambda} \leq \lambda_{k+1}/M^{k+1} \leq \lambda_k / M^k \leq \overline{\lambda};$$

**(vi)** $L(x, \cdot) : H \to H$ *is monotone over dom* $\varphi$ *for all* $x \in$ *dom* $\varphi$ *and Lipschitz continuous uniformly in* $x$ *with Lipschitz constant* $l > 0$ *over dom* $\varphi$*;*

**(vii)** *there exists* $\gamma > \overline{\lambda}/(2\beta')$ *such that, for all* $x, y \in$ *dom* $\varphi$*,*

*if* $\langle F(x), y - x \rangle + \varphi(y) - \varphi(x) \geq 0$ *holds, then*

$$\langle F(y) - L(y, y) + L(y, x), y - x \rangle + \varphi(y) - \varphi(x)$$

$$\geq \gamma \| (F(y) - L(y, y)) - (F(x) - L(y, x)) \|^2;$$

**(viii)** $\{\varphi^k\}_{k \in I\!N}, \varphi \in \Gamma_0(H)$ *are such that* $\varphi^k \stackrel{M}{\to} \varphi$ *and* $\varphi \leq \varphi^k$ *for all* $k$*. Moreover, for some solution* $x^*$ *of problem* $(GVIP)$*, there exist a constant* $\alpha > 1$ *and a sequence* $\{w^k\}_{k \in I\!N}$ *such that*

$$k^\alpha \| w^k - x^* \| \to 0 \quad \text{and} \quad k^\alpha | \varphi^k(w^k) - \varphi(x^*) | \to 0 \tag{4.5}$$

*i.e.,* $\{\varphi^k\}_{k \in I\!N}$ *converges to* $\varphi$ *at the order* $\alpha$ *at* $x^*$*.*

*Then, provided that $x^0 \in dom\ \varphi$, the sequence $\{x^k\}_{k\in I\!N}$ is bounded. Moreover,*

$$\lim_{k\to+\infty} \|x^{k+1} - x^k\| = 0 \text{ and}$$

$$\lim_{k\to+\infty} \|[F(x^k) - L(x^k, x^k)] - [F(x^*) - L(x^k, x^*)]\| = 0.$$

**Proof.** Let $x^*$ be the solution of problem $(GVIP)$ used in assumption $(viii)$. We consider the sequence of Lyapunov functions $\{\Gamma^k(x^*, \cdot)\}_{k\in I\!N}$ defined on $H$ by

$$\begin{aligned}\Gamma^k(x^*, x) &= (h^k(x^*) - h^k(x) - \langle\nabla h^k(x), x^* - x\rangle)/M^k \\ &\quad +(\lambda_k/M^k)(\langle F(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*)).\end{aligned} \tag{4.6}$$

From assumptions $(i)$, $(iii)$–$(v)$ and the fact that $x^*$ is a solution of $(GVIP)$, we derive immediately that, for all $x \in \text{dom } \varphi$ and all $k \in I\!N$,

$$\Gamma^k(x^*, x) \geq (\beta'/2)\|x - x^*\|^2 \quad \text{and} \quad \Gamma^{k+1}(x^*, x) \leq \Gamma^k(x^*, x). \tag{4.7}$$

Now, in order to prove that $\{x^k\}_{k\in I\!N}$ is bounded and that $\{x^{k+1} - x^k\}_{k\in I\!N}$ converges to 0, we will show that the following inequality holds for all $k \in I\!N$:

$$\begin{aligned}\alpha_k\Gamma^{k+1}(x^*, x^{k+1}) &\leq \Gamma^k(x^*, x^k) - c_1\|x^{k+1} - x^k\|^2 + T^k \\ &\quad -c_2\|F(x^k) - L(x^k, x^k) - F(x^*) + L(x^k, x^*)\|^2,\end{aligned} \tag{4.8}$$

where $c_1, c_2$ are positive constants, $\{\alpha_k\}_{k\in I\!N}$ is a sequence of positive numbers such that $0 < \alpha_0 \leq \alpha_k < 1$ for all $k$, $\lim_{k\to\infty} \alpha_k = 1$, $\lim_{k\to\infty} \prod_{j=0}^k \alpha_j^{-1} > 0$ and $\{T^k\}_{k\in I\!N}$ is a nonnegative sequence satisfying $\sum_{k=0}^{+\infty} T^k < +\infty$.

Using the second inequality of (4.7) and the definition of the Lyapunov function, we can write

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq \Gamma^k(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) = s_1 + s_2 + s_3, \tag{4.9}$$

with

$$\begin{aligned}s_1 &= (h^k(x^k) - h^k(x^{k+1}) + \langle\nabla h^k(x^k), x^{k+1} - x^k\rangle)/M^k, \\ s_2 &= \langle\nabla h^k(x^k) - \nabla h^k(x^{k+1}), x^* - x^{k+1}\rangle/M^k, \\ s_3 &= (\lambda_k/M^k)(\langle F(x^*), x^{k+1} - x^k\rangle + \varphi(x^{k+1}) - \varphi(x^k)).\end{aligned}$$

For $s_1$, we derive easily from assumptions $(i)$ and $(iv)$ that

$$s_1 \leq -(\beta'/2)\|x^{k+1} - x^k\|^2. \tag{4.10}$$

Now, using the sequence $\{w^k\}_{k\in I\!N}$ given in assumption $(viii)$, we can write $s_2$ as the sum of the two following terms:

$$\begin{aligned} s_{21} &= \langle \nabla h^k(x^k) - \nabla h^k(x^{k+1}), x^* - w^k\rangle / M^k, \\ s_{22} &= \langle \nabla h^k(x^k) - \nabla h^k(x^{k+1}), w^k - x^{k+1}\rangle / M^k. \end{aligned}$$

From assumptions $(ii)$ and $(iv)$, we deduce that

$$\begin{aligned} s_{21} &\leq \Lambda' \|x^{k+1} - x^k\| \|x^* - w^k\| \\ &\leq (\tau/2)\|x^{k+1} - x^k\|^2 + (1/2\tau)\Lambda'^2 \|x^* - w^k\|^2, \end{aligned} \tag{4.11}$$

where the second inequality holds for any $\tau > 0$ by Lemma 4.1.
Using the definition of $x^{k+1}$ with $x = w^k$ (see problem $(PAP^k)$), we obtain

$$\begin{aligned} s_{22} &\leq (\lambda_k/M^k)[\langle F(x^k) + L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle \\ &\quad + \varphi^k(w^k) - \varphi^k(x^{k+1})] \\ &= (\lambda_k/M^k)[\langle F(x^k), x^* - x^k\rangle + \varphi(x^*) - \varphi(x^k) \\ &\quad + \langle F(x^k), x^k - x^{k+1}\rangle + \langle F(x^k), w^k - x^*\rangle \\ &\quad + \varphi(x^k) - \varphi^k(x^{k+1}) + \varphi^k(w^k) - \varphi(x^*) \\ &\quad + \langle L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle]. \end{aligned} \tag{4.12}$$

By assumption $(vii)$ and since $\langle F(x^*), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*) \geq 0$, we get

$$\begin{aligned} &\langle F(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*) \\ &\qquad \geq \langle L(x^k, x^k) - L(x^k, x^*), x^k - x^*\rangle \\ &\qquad + \gamma\|(F(x^k) - L(x^k, x^k)) - (F(x^*) - L(x^k, x^*))\|^2. \end{aligned} \tag{4.13}$$

Gathering the fact that $\varphi \leq \varphi^k$ for all $k$ with (4.12), (4.13), using Lemma 4.1 and rearranging the terms, we obtain the following inequalities:

$$\begin{aligned} &s_{22} + s_3 \\ &\leq (\lambda_k/M^k)[\langle (F(x^k) - L(x^k, x^k)) - (F(x^*) - L(x^k, x^*)), x^k - x^{k+1}\rangle \\ &\quad + \langle (F(x^k) - L(x^k, x^k)) - (F(x^*) - L(x^k, x^*)), w^k - x^*\rangle \\ &\quad - \gamma\|(F(x^k) - L(x^k, x^k)) - (F(x^*) - L(x^k, x^*))\|^2 \\ &\quad + \langle F(x^*), w^k - x^*\rangle + \varphi^k(w^k) - \varphi(x^*) \\ &\quad + \langle L(x^k, x^{k+1}) - L(x^k, x^*), w^k - x^{k+1}\rangle] \end{aligned}$$

$$\leq (\lambda_k/M^k)[((\mu+\eta-2\gamma)/2)\|(F(x^k)-L(x^k,x^k))-(F(x^*)-L(x^k,x^*))\|^2$$
$$+(1/2\mu)\|x^{k+1}-x^k\|^2+(1/2\eta)\|w^k-x^*\|^2$$
$$+\|F(x^*)\|\|w^k-x^*\|+|\varphi^k(w^k)-\varphi(x^*)|$$
$$+\langle L(x^k,x^{k+1})-L(x^k,x^*),w^k-x^{k+1}\rangle], \tag{4.14}$$

where $\mu, \eta$ are any positive constants.
In order to treat the last term in the right–hand side of (4.14), we use successively assumption $(vi)$, Lemma (4.1) and the first inequality of (4.7) to get

$$\langle L(x^k,x^{k+1})-L(x^k,x^*),w^k-x^{k+1}\rangle \leq \langle L(x^k,x^{k+1})-L(x^k,x^*),w^k-x^*\rangle$$
$$\leq l\|x^{k+1}-x^*\|\|w^k-x^*\|$$
$$\leq (l\theta_k/2)\|x^{k+1}-x^*\|^2+(l/2\theta_k)\|w^k-x^*\|^2$$
$$\leq (l\theta_k/\beta')\Gamma^{k+1}(x^*,x^{k+1})+(l/2\theta_k)\|w^k-x^*\|^2, \tag{4.15}$$

where $\theta_k$ is any positive number.
Now, since $\gamma > \bar\lambda/(2\beta')$, we can choose $\mu$, $\eta$ and $\tau$ such that $\mu+\eta-2\gamma<0$ and $\tau+\bar\lambda\mu^{-1}<\beta'$. Moreover, we take $\theta_k=\theta(k+1)^{-\alpha}$ for all $k$, with $0<\theta<\beta'\bar\lambda^{-1}l^{-1}$.
Then gathering (4.9),(4.10),(4.11),(4.14),(4.15), using assumption $(v)$ and rearranging the terms, we obtain (4.8) with

$$\begin{aligned}
c_1 &= (1/2)(\beta'-\tau-\bar\lambda\mu^{-1}), \quad c_2 \;=\; (\underline{\lambda}/2)(2\gamma-\mu-\eta), \\
\alpha_k &= 1-l\bar\lambda\beta'^{-1}\theta_k, \qquad\qquad (4.16)\\
T^k &= (1/2)[\Lambda'^2\tau^{-1}+\bar\lambda\eta^{-1}+l\bar\lambda\theta_k^{-1}]\|x^*-w^k\|^2 \\
&\quad +\bar\lambda\|F(x^*)\|\|x^*-w^k\|+\bar\lambda|\varphi^k(w^k)-\varphi(x^*)|. \qquad (4.17)
\end{aligned}$$

With the choice made for $\mu$, $\eta$, $\tau$ and $\theta_k$, it is easy to see that the constants $c_1$ and $c_2$ are positive and that all the conditions required on the sequence $\{\alpha_k\}_{k\in I\!N}$ are satisfied. Moreover, it follows from condition $(viii)$ that the series $\sum_{k=0}^{+\infty}T^k$ is convergent. Dividing then each member of (4.8) by $\alpha_k$, we obtain the following inequality:

$$\Gamma^{k+1}(x^*,x^{k+1}) \leq \alpha_k^{-1}\Gamma^k(x^*,x^k)+\alpha_0^{-1}T^k, \quad \forall k\in I\!N.$$

Since $\lim_{k\to\infty}\prod_{j=0}^{k}\alpha_j^{-1} > 0$ and $\sum_{k=0}^{+\infty}\alpha_0^{-1}T^k$ is convergent, it follows from Lemma 4.2 that $\{\Gamma^k(x^*, x^k)\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Hence, it is convergent in $\mathbb{R}$. Using then (4.7), we deduce that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded and, passing to the limit in (4.8), that the sequences $\{\|x^{k+1} - x^k\|\}_{k\in\mathbb{N}}$ and $\{\|[F(x^k) - L(x^k, x^k)] - [F(x^*) - L(x^k, x^*)]\|\}_{k\in\mathbb{N}}$ converge to zero. □

We can now state the main convergence result.

**Proposition 4.3** *Let us suppose that all assumptions of Proposition 4.2 are fulfilled such that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded. The following conclusions can be derived:*

1. *If $F$ is weakly continuous on dom $\varphi$ and the functional $x \to \langle F(x), x\rangle$ is weakly lower semi–continuous on dom $\varphi$, then each weak limit point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a solution of problem $(GVIP)$.*

2. *If, in addition, assumption (viii) of Proposition 4.2 is satisfied for any solution of problem $(GVIP)$ with a common $\alpha > 1$ and the sequence of operators $\{\nabla h^k\}_{k\in\mathbb{N}}$ satisfies*

$$z^k \rightharpoonup z, z^k \in dom\ \varphi \Rightarrow \nabla h^k(z^k) - \nabla h^k(z) \rightharpoonup 0, \tag{4.18}$$

   *then the whole sequence $\{x^k\}_{k\in\mathbb{N}}$ weakly converges to some solution of problem $(GVIP)$.*

3. *If, moreover, $F$ is strongly monotone on dom $\varphi$, then $\{x^k\}_{k\in\mathbb{N}}$ strongly converges to the unique solution $x^*$ of problem $(GVIP)$.*

**Proof.** Conclusion 1 follows directly from Proposition 4.2 and Proposition 4.1.

To prove conclusion 2, we have to show that the sequence $\{x^k\}_{k\in\mathbb{N}}$ has a unique weak limit point. Assume that $\{x^k\}_{k\in\mathbb{N}}$ has two weak limit points $\bar{x}$ and $\tilde{x}$. By conclusion 1, these two points are solutions of problem $(GVIP)$ and from the proof of Proposition 4.2, the sequences $\{\Gamma^k(\tilde{x}, x^k)\}_{k\in\mathbb{N}}$ and $\{\Gamma^k(\bar{x}, x^k)\}_{k\in\mathbb{N}}$ are convergent. Let $\tilde{\Gamma}$ and $\bar{\Gamma}$ be their respective limits.
On the other hand, by definition of the Lyapunov function, we have, for all

$k \in I\!N$ and $x \in \mathrm{dom}\,\varphi$,

$$\begin{aligned}&\Gamma^k(\tilde{x}, x) - \Gamma^k(\bar{x}, x)\\ &\quad = (1/M^k)(h^k(\tilde{x}) - h^k(\bar{x}) - \langle \nabla h^k(x), \tilde{x} - \bar{x}\rangle)\\ &\qquad +(\lambda_k/M^k)(\langle F(\tilde{x}), \bar{x} - \tilde{x}\rangle + \varphi(\bar{x}) - \varphi(\tilde{x}) + \langle F(\tilde{x}) - F(\bar{x}), x - \bar{x}\rangle).\end{aligned}$$

Let $\{x^k\}_{k\in K\subset I\!N}$ be a subsequence of $\{x^k\}_{k\in I\!N}$ converging to $\bar{x}$. If we set $x = x^k$ in the above inequality, we can write:

$$\begin{aligned}&\Gamma^k(\tilde{x}, x^k) - \Gamma^k(\bar{x}, x^k)\\ &= (1/M^k)(h^k(\tilde{x}) - h^k(\bar{x}) - \langle \nabla h^k(x^k) - \nabla h^k(\bar{x}), \tilde{x} - \bar{x}\rangle - \langle \nabla h^k(\bar{x}), \tilde{x} - \bar{x}\rangle)\\ &\quad +(\lambda_k/M^k)(\langle F(\tilde{x}), \bar{x} - \tilde{x}\rangle + \varphi(\bar{x}) - \varphi(\tilde{x}) + \langle F(\tilde{x}) - F(\bar{x}), x^k - \bar{x}\rangle).\end{aligned}$$

From assumptions $(i)$, $(iv)$ of Proposition 4.2 and since $\tilde{x}$ is a solution of problem $(GVIP)$, we deduce that

$$\begin{aligned}\Gamma^k(\tilde{x}, x^k) - \Gamma^k(\bar{x}, x^k) \quad \geq \quad & (\beta'/2)\|\tilde{x} - \bar{x}\|^2\\ & -(1/M^k)\langle \nabla h^k(x^k) - \nabla h^k(\bar{x}), \tilde{x} - \bar{x}\rangle\\ & +(\lambda_k/M^k)\langle F(\tilde{x}) - F(\bar{x}), x^k - \bar{x}\rangle.\end{aligned}$$

Then, if we take the limit on $k \in K$, condition (4.18) on $\{\nabla h^k\}_{k\in I\!N}$ implies that

$$\tilde{\Gamma} - \bar{\Gamma} \geq (\beta'/2)\|\tilde{x} - \bar{x}\|^2.$$

Since the role of $\bar{x}$ and $\tilde{x}$ is symmetric, we also have that

$$\bar{\Gamma} - \tilde{\Gamma} \geq (\beta'/2)\|\bar{x} - \tilde{x}\|^2.$$

Gathering these two inequalities, we conclude that $\bar{x} = \tilde{x}$ which proves the uniqueness of the weak limit point for $\{x^k\}_{k\in I\!N}$.

Let $x^*$ denote this weak limit point. To obtain conclusion 3, we will show that when $F$ is strongly monotone on $\mathrm{dom}\,\varphi$, we also have that $\|x^k - x^*\| \to 0$. If we put together relations (4.9), (4.10), (4.11), (4.12) in the proof of

Proposition 4.2, we obtain that

$$\begin{aligned}
&\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \\
&\leq ((\tau - \beta')/2)\|x^{k+1} - x^k\|^2 + (\Lambda'^2/2\tau)\|x^* - w^k\|^2 \\
&\quad +(\lambda_k/M^k)[\langle F(x^k), x^* - x^k\rangle + \varphi(x^*) - \varphi(x^k) \\
&\quad +\langle F(x^k), x^k - x^{k+1}\rangle + \langle F(x^k), w^k - x^*\rangle \\
&\quad +\varphi(x^k) - \varphi^k(x^{k+1}) + \varphi^k(w^k) - \varphi(x^*) \\
&\quad +\langle L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle \\
&\quad +\langle F(x^*), x^{k+1} - x^k\rangle + \varphi(x^{k+1}) - \varphi(x^k)],
\end{aligned}$$

where $\tau$ is any positive constant.
When $F$ is strongly monotone on dom $\varphi$ (with constant $\bar{\alpha} > 0$), since $x^*$ is a solution of problem $(GVIP)$, we have that

$$\langle F(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*) \geq \bar{\alpha}\|x^k - x^*\|^2.$$

Using this and the fact that $\varphi \leq \varphi^k$ for all $k$, we deduce that

$$\begin{aligned}
&\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \\
&\leq ((\tau - \beta')/2)\|x^{k+1} - x^k\|^2 + (\Lambda'^2/2\tau)\|x^* - w^k\|^2 \\
&\quad +(\lambda_k/M^k)[-\bar{\alpha}\|x^k - x^*\|^2 + \varphi^k(w^k) - \varphi(x^*) \\
&\quad +\langle F(x^k), x^k - x^{k+1}\rangle + \langle F(x^k), w^k - x^*\rangle \\
&\quad +\langle L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle \\
&\quad +\langle F(x^*), x^{k+1} - x^k\rangle].
\end{aligned} \tag{4.19}$$

Let us now pass to the limit on $k$ in this inequality. From assumptions $(v)$,$(vi)$ and $(viii)$ of Proposition 4.2 and since we know that $\{\Gamma^k(x^*, x^k)\}_{k\in I\!N}$ is convergent, $\{\|x^{k+1} - x^k\|\}_{k\in I\!N}$ converges to zero and $F$ is weakly continuous, we conclude that $\{\|x^k - x^*\|\}_{k\in I\!N}$ converges to zero. This completes the proof. □

Observe that, for example, if $h^k(x) = (1/2)x^T x$ for all $x \in H$ and $k \in I\!N$, then $\nabla h$ is weakly continuous on $H$. Moreover, when $H$ is a finite dimensional space, $\nabla h$ is continuous in the strong topology and thus in the weak topology because the two topologies coincide.

**Remark 4.1** If we particularize Proposition 4.3 to the finite dimensional case, we obtain the (strong) convergence of the sequence $\{x^k\}_{k\in I\!N}$ to some solution of problem $(GVIP)$ provided that all assumptions Proposition 4.2 are satisfied, $F$ is continuous and assumption $(viii)$ of Proposition 4.2 is satisfied for any solution of the problem.

**Remark 4.2** Observe that if $L(x,\cdot)$ is monotone over dom $\varphi$ for all $x \in$ dom $\varphi$, and $F$ and $L$ satisfy assumption $(vii)$, then for any pair $x_1^*, x_2^*$ of solutions of problem $(GVIP)$, we have that

$$F(x_2^*) - F(x_1^*) = L(x_2^*, x_2^*) - L(x_2^*, x_1^*).$$

Indeed, by assumption $(vii)$ and since $\langle F(x_1^*), x_2^* - x_1^*\rangle + \varphi(x_2^*) - \varphi(x_1^*) \geq 0$, we have:

$$\langle F(x_2^*), x_2^* - x_1^*\rangle + \varphi(x_2^*) - \varphi(x_1^*) - \langle L(x_2^*, x_2^*) - L(x_2^*, x_1^*), x_2^* - x_1^*\rangle$$

$$\geq \gamma \|(F(x_2^*) - L(x_2^*, x_2^*)) - (F(x_1^*) - L(x_2^*, x_1^*))\|^2.$$

So, since $L(x_2^*,\cdot)$ is monotone and $x_2^*$ is a solution of problem $(GVIP)$, we deduce that

$$\|(F(x_2^*) - L(x_2^*, x_2^*)) - (F(x_1^*) - L(x_2^*, x_1^*))\| = 0.$$

In the particular case where $L = 0$ and thus $F$ satisfies the pseudo Dunn property, this relation amounts to say that the set

$$\{F(x^*) : x^* \text{ is a solution of problem } (GVIP)\}$$

is a singleton.

**Remark 4.3** When $\varphi^k = \varphi$ for all $k$, condition $(viii)$ of Proposition 4.2 is obviously satisfied by setting $w^k = x^*$ for all $k$. With this choice of $\{w^k\}_{k\in I\!N}$, it is unnecessary to require Ł$(x,\cdot)$ to be Lipschitz continuous over dom $\varphi$ uniformly in $x$ (see (4.15)).

**Remark 4.4** Assumption $(vii)$ of Proposition 4.2 can be replaced by the following slightly weaker condition:

**(vii')** *There exists* $\gamma > \overline{\lambda}/(2\beta')$ *and there exists a singlevalued operator* $\Delta : H \to H$ *such that:*

- $\Delta$ *is monotone over dom* $\varphi$*;*
- $L(x,\cdot) - \Delta$ *is monotone over dom* $\varphi$ *for all* $x \in$ *dom* $\varphi$*;*
- *for all* $x, y \in$ *dom* $\varphi$*,*

$$\text{if } \langle F(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0 \text{ holds, then}$$

$$\langle F(y) - (L(y,y) - \Delta(y)) + (L(y,x) - \Delta(x)), y - x\rangle + \varphi(y) - \varphi(x)$$

$$\geq \gamma \|(F(y) - L(y,y)) - (F(x) - L(y,x))\|^2.$$

The introduction of the operator $\Delta$ allows us to compare $L(x,\cdot)$ with a monotone mapping independent of the argument $x$. So, in the case where $L(x,\cdot)$ is monotone and does not depend on $x$, we can choose $\Delta(y) = L(x,y)$ for all $x, y \in \text{dom}\ \varphi$. With this choice of $\Delta$, condition $(vii')$ is similar to Condition $(RC)$ imposed by Renaud and Cohen (see Theorem 2.4). For the sake of completeness, let us state and justify the convergence result related to this modified condition $(vii')$.

***Proposition 4.4*** *Assume that the solution set of problem* $(GVIP)$ *is nonempty, that assumptions (i)–(vi) and (viii) of Proposition 4.2 and condition (vii') are satisfied.*

1. *Then, provided that* $x^0 \in$ *dom* $\varphi$*, the sequence* $\{x^k\}_{k\in \mathbb{N}}$ *generated by subproblems* $(PAP^k)$ *is bounded and such that* $\|x^{k+1} - x^k\| \to 0$*. Moreover, if* $F$ *is weakly continuous on dom* $\varphi$ *and the functional* $x \to \langle F(x), x\rangle$ *is weakly lower semi–continuous on dom* $\varphi$*, then each weak limit point of the sequence* $\{x^k\}_{k\in \mathbb{N}}$ *is a solution of problem* $(GVIP)$*.*

2. *If, in addition, assumption (viii) is satisfied for any solution of problem* $(GVIP)$ *with a common* $\alpha > 1$*, the operator* $\Delta$ *is weakly continuous and the sequence of operators* $\{\nabla h^k\}_{k\in \mathbb{N}}$ *satisfies (4.18), then the whole sequence* $\{x^k\}_{k\in \mathbb{N}}$ *weakly converges to some solution of problem* $(GVIP)$*.*

3. *If, moreover,* $F$ *is strongly monotone on dom* $\varphi$ *and* $L(x,\cdot) - \Delta$ *is Lipschitz continuous over dom* $\varphi$*, then* $\{x^k\}_{k\in \mathbb{N}}$ *strongly converges to the unique solution* $x^*$ *of problem* $(GVIP)$*.*

**Proof.** To show that conclusion 1 holds, we will review the proof of Proposition 4.2. So, let $x^*$ be the solution of problem $(GVIP)$ used in assumption $(viii)$ of Proposition 4.2. We add a nonnegative term to the Lyapounov function $\Gamma^k(x^*, \cdot)$ (see (4.6)) to define the sequence of functions $\{\Xi^k(x^*, \cdot)\}_{k\in I\!N}$:

$$\Xi^k(x^*, x) = \Gamma^k(x^*, x) + (\lambda_k/M^k)\langle \Delta(x^*) - \Delta(x), x^* - x\rangle.$$

In the same way as in the proof of Proposition 4.2, we can write

$$\Xi^{k+1}(x^*, x^{k+1}) - \Xi^k(x^*, x^k) \leq s_1 + s_2 + s_3 + s_4, \tag{4.20}$$

where $s_1, s_2, s_3$ are like in (4.9), and

$$s_4 = (\lambda_k/M^k)[\langle \Delta(x^*) - \Delta(x^{k+1}), x^* - x^{k+1}\rangle - \langle \Delta(x^*) - \Delta(x^k), x^* - x^k\rangle].$$

The terms $s_1$ and $s_2$ can be treated like in (4.10), (4.11) and (4.12). If we use condition $(vii')$ with $x = x^*$ and $y = x^k$ instead of (4.13) and we make the same manipulations as those used to get (4.14), we obtain:

$$\begin{aligned}
& s_{22} + s_3 + s_4 \\
& \leq (\lambda_k/M^k)[((\mu + \eta - 2\gamma)/2)\|(F(x^k) - L(x^k, x^k)) - (F(x^*) - L(x^k, x^*))\|^2 \\
& \quad +(1/2\mu)\|x^{k+1} - x^k\|^2 + (1/2\eta)\|w^k - x^*\|^2 \\
& \quad +\|F(x^*)\|\,\|w^k - x^*\| + |\varphi^k(w^k) - \varphi(x^*)| \\
& \quad +\langle L(x^k, x^{k+1}) - L(x^k, x^*), w^k - x^{k+1}\rangle + \langle \Delta(x^*) - \Delta(x^{k+1}), x^* - x^{k+1}\rangle],
\end{aligned}$$

where $\mu, \eta$ are any positive constants.
From the monotonicity of $L(x^k, \cdot) - \Delta$, we deduce that

$$\begin{aligned}
& \langle L(x^k, x^{k+1}) - L(x^k, x^*), w^k - x^{k+1}\rangle + \langle \Delta(x^*) - \Delta(x^{k+1}), x^* - x^{k+1}\rangle \\
& \leq \langle L(x^k, x^{k+1}) - L(x^k, x^*), w^k - x^*\rangle,
\end{aligned}$$

and this last term can be treated like in (4.15).
So, the same arguments as in the proof of Proposition 4.2 allow us to conclude that the sequence $\{\Xi^k(x^*, x^k)\}_{k\in I\!N}$ converges, $\{x^k\}_{k\in I\!N}$ is bounded and $\{\|x^{k+1} - x^k\|\}_{k\in I\!N}$ converges to zero. Moreover, if follows from Proposition 4.1 that any weak limit point of $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$.

To prove conclusion 2, let us first remark that for any $x, \tilde{x}, \bar{x} \in H$, we have

$$\begin{aligned}
&\Xi^k(\tilde{x}, x) - \Xi^k(\bar{x}, x) \\
&\quad = \Gamma^k(\tilde{x}, x) - \Gamma^k(\bar{x}, x) \\
&\qquad +(\lambda_k/M^k)[\langle \Delta(\tilde{x}) - \Delta(x), \tilde{x} - x\rangle - \langle \Delta(\bar{x}) - \Delta(x), \bar{x} - x\rangle] \\
&\quad = \Gamma^k(\tilde{x}, x) - \Gamma^k(\bar{x}, x) \\
&\qquad +(\lambda_k/M^k)[\langle \Delta(\tilde{x}) - \Delta(\bar{x}), \tilde{x} - x\rangle + \langle \Delta(\bar{x}) - \Delta(x), \tilde{x} - \bar{x}\rangle].
\end{aligned}$$

Hence, by using the same reasoning as in the proof of part 2 of Proposition 4.3, the weak continuity and the monotonicity of $\Delta$, we derive that the sequence $\{x^k\}_{k\in I\!N}$ has only one weak limit point.

If we denote by $x^*$ this weak limit point and if we assume the strong monotonicity of $F$, we can use relation (4.19) obtained in the proof of part 3 of Proposition 4.3 to obtain that

$$\begin{aligned}
&\Xi^{k+1}(x^*, x^{k+1}) - \Xi^k(x^*, x^k) \\
&\quad \leq ((\tau - \beta')/2)\|x^{k+1} - x^k\|^2 + (\Lambda'^2/2\tau)\|x^* - w^k\|^2 \\
&\qquad +(\lambda_k/M^k)[-\bar{\alpha}\|x^k - x^*\|^2 + \varphi^k(w^k) - \varphi(x^*) \\
&\qquad +\langle F(x^k), x^k - x^{k+1}\rangle + \langle F(x^k), w^k - x^*\rangle + \langle F(x^*), x^{k+1} - x^k\rangle \\
&\qquad +\langle L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle \\
&\qquad +\langle \Delta(x^*) - \Delta(x^{k+1}), x^* - x^{k+1}\rangle - \langle \Delta(x^*) - \Delta(x^k), x^* - x^k\rangle],
\end{aligned}$$

where $\tau$ is any positive constant.
By observing that

$$\begin{aligned}
&\langle L(x^k, x^{k+1}) - L(x^k, x^k), w^k - x^{k+1}\rangle + \langle \Delta(x^*) - \Delta(x^{k+1}), x^* - x^{k+1}\rangle \\
&\quad -\langle \Delta(x^*) - \Delta(x^k), x^* - x^k\rangle \\
&= \langle [L(x^k, x^{k+1}) - \Delta(x^{k+1})] - [L(x^k, x^k) - \Delta(x^k)], w^k - x^{k+1}\rangle \\
&\quad +\langle \Delta(x^{k+1}) - \Delta(x^k), w^k - x^*\rangle + \langle \Delta(x^*) - \Delta(x^k), x^k - x^{k+1}\rangle,
\end{aligned}$$

we obtain

$$\Xi^{k+1}(x^*, x^{k+1}) - \Xi^k(x^*, x^k)$$
$$\leq ((\tau - \beta')/2)\|x^{k+1} - x^k\|^2 + (\Lambda'^2/2\tau)\|x^* - w^k\|^2$$
$$+(\lambda_k/M^k)[-\bar{\alpha}\|x^k - x^*\|^2 + \varphi^k(w^k) - \varphi(x^*)$$
$$+\langle F(x^k), x^k - x^{k+1}\rangle + \langle F(x^k), w^k - x^*\rangle + \langle F(x^*), x^{k+1} - x^k\rangle$$
$$+\|[L(x^k, x^{k+1}) - \Delta(x^{k+1})] - [L(x^k, x^k) - \Delta(x^k)]\| \, \|w^k - x^{k+1}\|$$
$$+\langle \Delta(x^{k+1}) - \Delta(x^k), w^k - x^*\rangle + \langle \Delta(x^*) - \Delta(x^k), x^k - x^{k+1}\rangle].$$

Let us now pass to the limit in this inequality. From assumptions $(v)$ and $(viii)$ of Proposition 4.2, since $L(x^k, \cdot) - \Delta$ is Lipschitz continuous, $F$ and $\Delta$ are weakly continuous, $\{\Xi^k(x^*, x^k)\}_{k\in I\!N}$ converges and $\{\|x^{k+1} - x^k\|\}_{k\in I\!N}$ converges to zero, we conclude that $\{\|x^k - x^*\|\}_{k\in I\!N}$ converges to zero. $\square$

**Remark 4.5** If we focus somewhat on assumptions $(vi)$ and $(vii)$ of Proposition 4.2, it becomes clear that our analysis can be applied to decomposition of variational inequalitites with a mapping $F$ which needs not necessarily to satisfy the pseudo Dunn property. More precisely, assume that $F$ can be expressed as the sum of two mappings $F_1$ and $F_2$. If we set $L(x, y) = L'(x, y) + F_2(y)$ in the perturbed scheme $(PAP^k)$, $x^{k+1}$ is then characterized as the solution of the following subproblem:

$$(PPAP^k) \begin{cases} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F_1(x^k) + F_2(x^{k+1}) + L'(x^k, x^{k+1}) - L'(x^k, x^k) \\ +\lambda_k^{-1}(\nabla h^k(x^{k+1}) - \nabla h^k(x^k)), x - x^{k+1}\rangle \\ +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

We see that $F_1$ is fixed at the current iterate $x^k$ while $F_2$ is considered at a variable point. In that case, assumption $(vi)$ requires that $L'(x, \cdot) + F_2$ is monotone over dom $\varphi$ for all $x \in H$ and Lipschitz continuous uniformly in $x$. On the other hand, assumption $(vii)$ reduces to

**(vii")** *there exists $\gamma > \overline{\lambda}/(2\beta')$ such that for all $x, y \in$ dom $\varphi$,* *if* $\langle F(x), y-x\rangle + \varphi(y) - \varphi(x) \geq 0$ *holds, then*

$$\langle F_1(y) + F_2(x) - L'(y,y) + L'(y,x), y-x\rangle + \varphi(y) - \varphi(x)$$

$$\geq \gamma \|(F_1(y) - L'(y,y)) - (F_1(x) - L'(y,x))\|^2.$$

Roughly speaking, it amounts to require that $F_2$ be monotone and Lipschitz continuous and that $F_1$ be linked with $L'$ by a kind of pseudo Dunn property. In the extreme situation where we take $L' = 0, F = F_2\,(F_1 = 0)$ such that property $(vii'')$ is trivially true, we recover a perturbed and generalized instance of the proximal point algorithm:

$$\begin{cases} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F(x^{k+1}) + \lambda_k^{-1}(\nabla h^k(x^{k+1}) - \nabla h^k(x^k)), x - x^{k+1}\rangle \\ +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

This shows that the scheme characterized by subproblems $(PPAP^k)$ combines the auxiliary problem principle and the proximal point procedure. This is the reason why Kaplan and Tichatschke call it the proximal auxiliary problem method (see [66]). In that recent paper, they generalize our analysis in the sense that they obtain a convergence result for the scheme $(PPAP^k)$ under assumption $(vii'')$ where $F_2$ can be multivalued and the strong convexity of the functions $\{h^k\}_{k\in I\!N}$ is replaced by a condition of strong monotonicity on the operators $F_2 + \lambda_k^{-1}\nabla h^k$.

**Remark 4.6** When $\varphi^k$ plays the role of a barrier function associated with the constrained set $C$ in problem $(VIP)$, our scheme generates a sequence of points that lie in the interior of $C$. However, the philosophy is quite different from that of interior point path-following algorithms for solving variational inequalities (see, for example, [94], [116], [130], [131]).
Consider that, at each barrier parameter $\nu > 0$ is associated the following barrier subproblem:

$$(BP_\nu)\begin{cases} \text{find } x(\nu) \in \text{int}(C) \text{ such that, for all } x \in C, \\ \langle F(x(\nu)) + \nabla b(\nu, x(\nu)), x - x(\nu)\rangle \geq 0, \end{cases}$$

where $b$ denotes the barrier function preventing from going out of the set $C$. The sequence $\{x(\nu)\}_{\nu>0}$ generated by this way constitutes what is called

the central path.
At a current interior point $x^k$ and for a fixed $\nu_k$, path-following methods solve approximately subproblem $(BP_{\nu_k})$ by performing Newton steps to obtain some point $x^{k+1}$ sufficiently close to $x(\nu_k)$. In our scheme, the approximation $x^{k+1}$ of $x(\nu_k)$ is obtained by making one iteration of the auxiliary problem method applied to the operator $F$, i.e. by solving:

$$\langle F(x^k) + \lambda_k^{-1}(\Omega^k(x^{k+1}) - \Omega^k(x^k)) + \nabla b(\nu_k, x^{k+1}), x - x^{k+1}\rangle \geq 0, \forall x \in C.$$

## 4.2 Convergence Results for Particular Choices of the Auxiliary Operators

Our purpose here is to derive from Proposition 4.3 generalizations of well–known convergence results in the nonperturbed setting (see Chapter 2 Section 2.2.1). These different results will be obtained by choosing adequately $h^k$ and $L$ in the decomposition of the auxiliary operator $\Omega$ (see (4.4)).

### 4.2.1 The Symmetric Case

Let us first consider the symmetric case where, for all $k \in I\!N$, $\Omega(x^k, \cdot) = \Omega^k = \nabla K^k$ with $K^k$ some continuously differentiable and strongly convex function on dom $\varphi$. In this case, problem $(PAP^k)$ reduces to the optimization problem $(PSAP^k)$. If we choose $h^k = K^k$ for all $k$ and $L = 0$ in (4.4), assumption $(vi)$ of Proposition 4.2 obviously holds while assumption $(vii)$ of the same theorem amounts to express that $F$ has the $\varphi$–pseudo Dunn property on dom $\varphi$. Moreover, it follows from Remark 4.2 that any pair $x_1^*, x_2^*$ of solutions of problem $(GVIP)$ is such that $F(x_1^*) = F(x_2^*)$. This common value will be denoted by $F^*$ in the sequel. So, we derive from Proposition 4.3 the following convergence result related to subproblems $(PSAP^k)$:

**Proposition 4.5** *Suppose that the solution set of problem $(GVIP)$ is nonempty and that assumptions (i)–(v) and (viii) of Proposition 4.2 are satisfied with $h^k$ replaced by $K^k$ for all $k \in I\!N$. Assume, in addition, that the mapping $F$ has the $\varphi$–pseudo Dunn property with modulus $\gamma > \overline{\lambda}/(2\beta')$ on dom $\varphi$.
Then, provided that $x^0 \in$ dom $\varphi$, the sequence $\{x^k\}_{k\in I\!N}$ associated with problems $(PSAP^k)$, $k \in I\!N$, is bounded, such that $\{\|x^{k+1} - x^k\|\}_{k\in I\!N} \to 0$ and the sequence $\{F(x^k)\}_{k\in I\!N}$ converges to $F^*$. Moreover the same conclusions as in Proposition 4.3 can be deduced with $h^k$ replaced by $K^k$ for all $k$.* □

In comparison with Theorem 3.2, we do not need the mapping $F$ to be strongly monotone and Lipschitz continuous but only satisfy the $\varphi$–pseudo Dunn property. This enlarges the class of problems that can be considered, for example, by allowing multiple solutions. But, in return, we have to impose a speed of convergence on the sequence $\{\varphi^k\}_{k\in I\!N}$ described by assumption $(viii)$ of Proposition 4.2 while they only require the Mosco-convergence.

On the other hand, without perturbation on $\varphi$, with $K^k = K$ for all $k \in I\!N$ and $H = I\!R^n$, we recover Theorem 2.1 as a special case of Proposition 4.5.

From Proposition 4.5, we can derive new convergence results for the perturbed symmetric projection method. Recall that this method is devoted to solve problem $(VIP)$ and is characterized by the fact that $K^k(x) = (1/2)x^T D^k x$ for all $x \in I\!R^n$ and $k \in I\!N$, where $\{D^k\}_{k\in I\!N}$ is some sequence of symmetric positive definite matrices. To introduce perturbations, we consider a sequence $\{C^k\}_{k\in I\!N}$ of nonempty closed convex subsets of $C$ and we take for $\{\varphi^k\}_{k\in I\!N}$ the corresponding indicator functions. Then, problem $(PAP^k)$, with $\lambda_k = 1$ for all $k$, can be expressed as follows:

find $x^{k+1} \in C^k$ such that $\langle F(x^k)+D^k(x^{k+1}-x^k), x-x^{k+1}\rangle \geq 0, \quad \forall x \in C^k.$

It is easy to see that $x^{k+1}$ is precisely the projection with respect to the $D^k$–norm of the point $x^k - (D^k)^{-1}F(x^k)$ onto the closed convex set $C^k$.

Observe then that assumptions $(i)$–$(v)$ of Proposition 4.2 only concern the sequence $\{K^k\}_{k\in I\!N}$ and thus the sequence of matrices $\{D^k\}_{k\in I\!N}$. The next conditions on this sequence are sufficient to get these assumptions:

**(C1)** there exist two constants $\beta$ and $\Lambda$ such that

$$0 < \beta \leq \lambda_{min}(D^k) \leq \lambda_{max}(D^k) \leq \Lambda, \qquad \forall k \in I\!N;$$

**(C2)** for each $k \in I\!N$, there exists $\eta_k \geq 1$ satisfying, for all $x, y \in C$,

$$\|x-y\|^2_{D^{k+1}} \leq \eta_k \|x-y\|^2_{D^k};$$

**(C3)** there exist two positive constants $\underline{M}$ and $\overline{M}$ such that

$$0 < \underline{M} \leq M^k \leq \overline{M}, \qquad \forall k \in I\!N,$$

where $M^k = \prod_{j=0}^{k-1} \eta_j$.

For example, if for all $k \in I\!N$, $D^k = \zeta_k D$ where $D$ is a symmetric positive definite matrix and the sequence $\{\zeta_k\}_{k\in I\!N}$ is such that $0 < \underline{\zeta} \leq \zeta_k \leq \zeta_{k+1} \leq \overline{\zeta}$ for all $k$, then the sequence $\{D^k\}_{k\in I\!N}$ satisfies conditions (C1)–(C3) above. Another example where these conditions hold is given by a sequence of matrices $\{D^k\}_{k\in I\!N}$ which satisfies (C1) and $\|D^{k+1} - D^k\| \leq \theta_k$ for all $k$ where $\{\theta_k\}_{k\in I\!N}$ is a sequence of positive numbers such that $\sum_{k=0}^{+\infty} \theta_k < +\infty$.

The corresponding convergence result can then be stated as follows:

**Theorem 4.1** ($H = I\!R^n$) *Suppose that the solution set of problem* ($VIP$) *is nonempty and that the mapping* $F$ *is continuous and has the* $\varphi$*–pseudo Dunn property with modulus* $\gamma > 0$ *on* $C$. *If* $\{D^k\}_{k\in I\!N}$ *satisfies conditions* $(C1) - (C3)$ *above, if* $\overline{M} < 2\gamma\beta\underline{M}$ *and if assumption (viii) of Proposition 4.2 is satisfied for any solution of problem* ($VIP$)*, then, provided that* $x^0 \in$ *dom* $\varphi$*, the sequence* $\{x^k\}_{k\in I\!N}$ *generated by the perturbed symmetric projection method is convergent to a solution of problem* ($VIP$). □

When $C^k = C$ and $D^k = D$ for all $k \in I\!N$, this theorem extends a convergence result obtained by Marcotte and Wu ([83], Theorem 2.1).

In the very particular situation where $F = 0$, problem ($GVIP$) reduces to minimize the function $\varphi$ on $H$. If we choose $K^k(x) = (1/2)\|x\|^2$ for all $x$ and $k$, it is easy to see that the scheme based on problems ($PSAP^k$), $k \in I\!N$, is nothing else than the proximal point algorithm:

$$\text{find } x^{k+1} \in \text{ argmin } \{\varphi^k(x) + (1/2\lambda_k)\|x - x^k\|^2)\}.$$

In that case, we obtain the following theorem:

**Theorem 4.2** *If* $\varphi$ *admits at least a minimum, if assumption (viii) of Proposition 4.2 is satisfied for any minimum of* $\varphi$ *and if* $0 < \underline{\lambda} \leq \lambda_{k+1} \leq \lambda_k \leq \overline{\lambda}$ *for all* $k$*, then, provided that* $x^0 \in$ *dom* $\varphi$*, the sequence* $\{x^k\}_{k\in I\!N}$ *generated by the perturbed proximal point algorithm weakly converges to a minimum of* $\varphi$. □

The convergence of this algorithm has been analyzed in [65] in the case where the sequence $\{\varphi^k\}_{k\in I\!N}$ uniformly converges to $\varphi$, and in [73], [125] under different assumptions that refer to the notion of variational semi–distance between $\varphi^k$ and $\varphi$ introduced by Attouch and Wets in [7] (see Definition 3.3).

**Remark 4.7** By naturally choosing $h^k = K^k$ for all $k$ and $L = 0$ in (4.4) to obtain Proposition 4.5, we require $F$ to satisfy the $\varphi$–pseudo Dunn property. This condition is less flexible than condition $(vii)$ of Proposition 4.2 which allows to consider other sequences of functions $\{h^k\}_{k\in I\!N}$, the only constraint being that $F$ and $L(x^k,\cdot) = \lambda_k^{-1}(\nabla K^k - \nabla h^k)$ satisfy assumption $(vii)$ for all $k$. This is illustrated in the following example in $I\!R^2$ where:

- $\varphi$ is the indicator function of $C = I\!R \times \{0\} \subset I\!R^2$;
- $F(x) = Qx \;\text{ with }\; Q = \begin{pmatrix} 2 & -2 \\ 2 & 0 \end{pmatrix}, \quad \forall x \in C;$
- $K^k(x) = K(x) = (1/2)x^T Dx \;\text{ with }\; D = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \forall x, \quad \forall k;$
- $\lambda_k = 1, \quad \forall k \in I\!N.$

In this case, $\nabla K$ is Lipschitz continuous with modulus $\sqrt{5}$ on $C$ and strongly monotone with modulus 2 on $C$. In addition, $F$ does not satisfy the $\varphi$–pseudo Dunn property with modulus $\gamma > 1/4$ on $C$ so that Proposition 4.5 cannot be applied. However, if we choose $h^k(x) = h(x) = x^T x$ for all $x$ and all $k$, it can be easily seen that $\nabla h$ is Lipshitz continuous and strongly monotone with modulus 2 on $C$ and that $L(x,\cdot) = \nabla K - \nabla h$ is monotone and Lipschitz continuous on $C$. Moreover, little calculus shows that assumption $(vii)$ of Proposition 4.2 linking $F$ with $L$ is satisfied for any $\gamma$ such that $1/4 < \gamma \leq 2/5$. So, for this example, Proposition 4.3 alone can ensure the convergence of the process.

### 4.2.2 The Nonsymmetric Case

Now, we turn to the more general case of nonsymmetric auxiliary operators. We will derive from Proposition 4.3 generalizations of some results due to Tseng (see [126]) and to Pang and Chan (see [99]).
Assume that $\lambda_k = 1$ for all $k$ and that there exists a symmetric positive definite matrix $G$ such that all the operators $\Omega(x^k,\cdot) - G,\ k \in I\!N$, are monotone. Then, we can choose $h^k$ and $L$ in (4.4) such that:

$$h^k(y) = (1/2)y^T Gy \quad \text{and} \quad L(x^k, y) = \Omega(x^k, y) - Gy, \quad \forall y, \quad \forall k.$$

In this special case, we observe easily that assumptions $(i)$–$(v)$ of Proposition 4.2 hold with $\beta_k = \beta = \beta' = \lambda_{min}(G)$, $\Lambda_k = \Lambda = \Lambda' = \lambda_{max}(G)$,

$M^k = \eta_k = 1$ for all $k$. Moreover, assumption $(vi)$ is also immediately satisfied if, for all $x \in \text{dom}\,\varphi$, $\Omega(x,\cdot) - G$ is monotone over dom $\varphi$ for all $x \in \text{dom}\,\varphi$ and $\Omega(x,\cdot)$ is Lipschitz continuous over dom $\varphi$ uniformly in $x$. For the sake of convenience, we introduce the following condition:

**Condition (S):**
*there exists a symmetric positive definite matrix $G$ such that, for all $x \in$ dom $\varphi$, $\Omega(x,\cdot) - G$ is monotone over dom $\varphi$ and there exists a constant $\gamma > (2\lambda_{min}(G))^{-1}$ such that, for all $x, y \in$ dom $\varphi$,*

$$\text{if } \langle F(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0 \text{ holds, then}$$

$$\langle F(y) - (\Omega(y,y) - Gy) + (\Omega(y,x) - Gx), y - x\rangle + \varphi(y) - \varphi(x)$$
$$\geq \gamma \|F(y) - (\Omega(y,y) - Gy) - F(x) + (\Omega(y,x) - Gx)\|^2.$$

The following convergence result related to subproblems $(PAP^k)$, $k \in I\!N$, can then be deduced from Proposition 4.3:

**Theorem 4.3** *Assume that the solution set of problem $(GVIP)$ is nonempty and that the following conditions hold:*

**(i)** *$\Omega(x,\cdot) : H \to H$ is Lipschitz continuous over dom $\varphi$ uniformly in $x$;*

**(ii)** *$F$ and $\Omega$ satisfy Condition $(S)$;*

**(iii)** *$\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ are such that assumption (viii) of Proposition 4.2 holds.*

*Then, provided that $x^0 \in$ dom $\varphi$, the sequence $\{x^k\}_{k\in I\!N}$ associated with problems $(PAP^k)$, $k \in I\!N$, is bounded. Moreover, the same conclusions as in Proposition 4.3 can be derived (Note that in this case, condition (4.18) is immediately satisfied).* □

When for all $x, y \in H$, $\Omega(x,y) = D(x)y$ with $D(x)$ an $(n \times n)$ positive definite matrix and $\varphi^k = \varphi$ for all $k$, Pang and Chan ([99]) propose a convergence result using, instead of Condition $(S)$, a contraction condition that links $F$ and $\Omega$ (see Condition $(PCH)$ in Theorem 2.2). In our more general setting, this condition can be expressed as:

**Condition (GPCH):**

*there exists a symmetric positive definite matrix $G$ such that, for all $x \in$ dom $\varphi$, $\Omega(x, \cdot) - G$ is monotone over dom $\varphi$ and there exists a positive constant $b < 1$ such that, for all $x, y \in$ dom $\varphi$:*

$$\|G^{-1}[F(y) - F(x) - (\Omega(y, y) - \Omega(y, x))]\|_G \leq b\|y - x\|_G.$$

Our purpose now is to derive from Theorem 4.3 the following result which is an extension of Theorem 2.2:

**Theorem 4.4** *Assume that the following conditions hold:*

**(i)** $\Omega(x, \cdot) : H \to H$ *is Lipschitz continuous over dom $\varphi$ uniformly in $x$;*

**(ii)** $F$ *and* $\Omega$ *satisfy Condition (GPCH);*

**(iii)** $\{\varphi^k\}_{k \in I\!N}, \varphi \in \Gamma_0(H)$ *are such that assumption (viii) of Proposition 4.2 holds.*

*Then, provided that $x^0 \in$ dom $\varphi$, $F$ is weakly continuous on dom $\varphi$ and the functional $x \to \langle F(x), x \rangle$ is weakly lower semi–continuous on dom $\varphi$, then the sequence $\{x^k\}_{k \in I\!N}$ generated by problems $(PAP^k)$ strongly converges to the unique solution of problem $(GVIP)$.*

**Proof.** First, we will prove that, without loss of generality, we can suppose that $G = I$ in Condition $(GPCH)$. For that, we follow the same kind of procedure as Tseng in [126]: we introduce a scaling on the variables in such a way that, in the transformed space, the corresponding Condition $(GPCH)$ is satisfied with $G = I$. More precisely, we consider the linear transformation given by $\widehat{x} = G^{1/2}x$. With this scaling, the new functions corresponding to $F, \varphi, \varphi^k$ and $\Omega$ will be denoted by $\widehat{F}, \widehat{\varphi}, \widehat{\varphi^k}$ and $\widehat{\Omega}$ respectively and defined, for all $x, y \in H$, by

$$\begin{aligned}
\widehat{F}(\widehat{x}) &= G^{-1/2}F(G^{-1/2}\widehat{x}), \\
\widehat{\varphi}(\widehat{x}) &= \varphi(G^{-1/2}\widehat{x}) \quad \text{and} \quad \widehat{\varphi^k}(\widehat{x}) = \varphi^k(G^{-1/2}\widehat{x}), \\
\widehat{\Omega}(\widehat{x}, \widehat{y}) &= G^{-1/2}\Omega(G^{-1/2}\widehat{x}, G^{-1/2}\widehat{y}).
\end{aligned}$$

According to this scaling, problem $(GVIP)$ and problem $(PAP^k)$ are respectively equivalent to the problems:

$(\widehat{GVIP})$ find $\widehat{x^*} \in H$ such that $\langle \widehat{F}(\widehat{x^*}), \widehat{x} - \widehat{x^*} \rangle + \widehat{\varphi}(\widehat{x}) - \widehat{\varphi}(\widehat{x^*}) \geq 0$, $\forall \widehat{x} \in H$, and

$$(\widehat{PAP^k})\begin{cases}\text{find } \widehat{x^{k+1}} \in H \text{ such that, for all } \widehat{x} \in H,\\[2mm] \langle \widehat{F}(\widehat{x^k}) + \widehat{\Omega}(\widehat{x^k}, \widehat{x^{k+1}}) - \widehat{\Omega}(\widehat{x^k}, \widehat{x^k}), \widehat{x} - \widehat{x^{k+1}}\rangle \\[2mm] \qquad\qquad\qquad + \widehat{\varphi^k}(\widehat{x}) - \widehat{\varphi^k}(\widehat{x^{k+1}}) \geq 0,\end{cases}$$

in the sense that if $x^*$ (resp. $x^{k+1}$) is some solution of problem $(GVIP)$ (resp. $(PAP^k)$), then $\widehat{x^*}$ (resp. $\widehat{x^{k+1}}$) is a solution of problem $(\widehat{GVIP})$ (resp. $(\widehat{PAP^k})$) and reciprocally. So, it is sufficient to prove that problem $(\widehat{GVIP})$ admits a unique solution and that the sequence $\{\widehat{x^k}\}_{k\in I\!N}$ strongly converges to this solution to obtain the desired result.
Since $\Omega(x,\cdot) - G$ is monotone on dom $\varphi$ for all $x \in$ dom $\varphi$ if and only if $\widehat{\Omega}(\widehat{x},\cdot) - I$ is monotone on dom $\widehat{\varphi}$ for all $\widehat{x} \in$ dom $\widehat{\varphi}$, we have that Condition $(GPCH)$ is equivalent to:

**Condition(G$\widehat{\textbf{PC}}$H):**
*For all $\widehat{x} \in$ dom $\widehat{\varphi}$, $\widehat{\Omega}(\widehat{x},\cdot) - I$ is monotone over dom $\widehat{\varphi}$ and there exists a positive constant $b < 1$ such that, for all $\widehat{x}, \widehat{y} \in$ dom $\widehat{\varphi}$:*

$$\|\widehat{F}(\widehat{y}) - \widehat{F}(\widehat{x}) - (\widehat{\Omega}(\widehat{y},\widehat{y}) - \widehat{\Omega}(\widehat{y},\widehat{x}))\| \leq b\|\widehat{y} - \widehat{x}\|.$$

So, since assumptions $(i)$ and $(iii)$ with $\varphi, \varphi^k$ and $\Omega$ replaced respectively by $\widehat{\varphi}, \widehat{\varphi^k}$ and $\widehat{\Omega}$, are satisfied, we can, without loss of generality, suppose that $G = I$ in Condition $(GPCH)$.
We prove now that Condition $(GPCH)$ with $G = I$ implies the following condition:

**Condition (SI):**
*For all $x \in$ dom $\varphi$, $\Omega(x,\cdot) - I$ is monotone over dom $\varphi$ and there exists a constant $\gamma > 1/2$ such that, for all $x, y \in$ dom $\varphi$,*

$$\text{if } \langle F(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0 \text{ holds, then}$$

$$\langle F(y) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle + \varphi(y) - \varphi(x)$$
$$\geq \gamma\|F(y) - (\Omega(y,y) - y) - F(x) + (\Omega(y,x) - x)\|^2.$$

Let $x, y \in$ dom $\varphi$. If $F$ and $\Omega$ satisfy Condition $(GPCH)$ with $G = I$, we have

$$\begin{aligned}
&\|F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x)\|^2 \\
&\quad = \|F(y) - F(x) - \Omega(y,y) + \Omega(y,x)\|^2 + \|y - x\|^2 \\
&\qquad +2\langle F(y) - F(x) - \Omega(y,y) + \Omega(y,x), y - x\rangle \\
&\quad \le 2\langle F(y) - F(x) - \Omega(y,y) + \Omega(y,x), y - x\rangle + (b^2+1)\|y - x\|^2 \\
&\quad = 2\langle F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle \\
&\qquad +(b^2-1)\|y - x\|^2.
\end{aligned}$$

Or equivalently,

$$\begin{aligned}
&2\langle F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle \ge (1-b^2)\|y - x\|^2 \\
&+\|F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x)\|^2.
\end{aligned} \tag{4.21}$$

Moreover, by using again Condition $(GPCH)$ with $G = I$, we obtain

$$\|F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x)\| \le (b+1)\|y - x\|. \tag{4.22}$$

So, incorporating (4.22) in (4.21), we have

$$\begin{aligned}
&\langle F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle \\
&\quad \ge \gamma\|F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x)\|^2,
\end{aligned}$$

where $\gamma = 1/2 + (1-b^2)/(2(1+b)^2)$. Since $\gamma > 1/2$, we deduce that Condition $(SI)$ holds. Observing then that Condition $(S)$ is satisfied with $G = I$, we can apply Theorem 4.3. Since in this case $\{\nabla h^k\}_{k\in I\!N}$ satisfies condition (4.18), to conclude that the sequence $\{x^k\}_{k\in I\!N}$ strongly converges to the unique solution of problem $(GVIP)$, it remains to prove that $F$ is strongly monotone when Condition $(GPCH)$ with $G = I$ is satisfied.
Let $x, y \in \mathrm{dom}\ \varphi$. From (4.21), we deduce that

$$\begin{aligned}
&\langle F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle \\
&\ge (1-b^2)\|y - x\|^2/2.
\end{aligned} \tag{4.23}$$

On the other hand, $\Omega(y,\cdot) - I$ being monotone on dom $\varphi$, it follows that

$$\begin{aligned}
&\langle F(y) - F(x) - (\Omega(y,y) - y) + (\Omega(y,x) - x), y - x\rangle \\
&\le \langle F(y) - F(x), y - x\rangle.
\end{aligned} \tag{4.24}$$

Gathering (4.23) and (4.24), we get

$$\langle F(y) - F(x), y - x \rangle \geq (1 - b^2)\|y - x\|^2/2.$$

This means that $F$ is strongly monotone over dom $\varphi$ and the proof is complete. $\square$

The following proposition will be used to derive corollaries from Theorem 4.4.

**Proposition 4.6** *A sufficient condition for (GPCH) to hold is the following:*

**Condition (GPCHbis):**
*there exists a symmetric positive definite matrix $G$ such that, for all $x \in$ dom $\varphi$, $\Omega(x, \cdot) - G$ is monotone over dom $\varphi$ and there exists a positive constant $b' < \lambda_{min}(G)$ such that, for all $x, y \in$ dom $\varphi$:*

$$\|F(y) - F(x) - (\Omega(y, y) - \Omega(y, x))\| \leq b'\|y - x\|.$$

**Proof.** Recall that, for all $z \in H$,

$$\lambda_{min}(G)\|z\|^2 \leq \|z\|_G^2 = z^T G z \leq \lambda_{max}(G)\|z\|^2.$$

If we use the notation

$$R(x, y) = F(y) - F(x) - (\Omega(y, y) - \Omega(y, x)),$$

and we suppose that Condition ($GPCHbis$) is satisfied, the following inequalities can be deduced:

$$\begin{array}{rcl} \|G^{-1}R(x, y)\|_G^2 & = & \|R(x, y)\|_{G^{-1}}^2 \\ & \leq & \lambda_{max}(G^{-1})\,\|R(x, y)\|^2 \\ & \leq & (\lambda_{min}(G))^{-1}b'^2\,\|y - x\|^2 \\ & \leq & (\lambda_{min}(G))^{-2}b'^2\,\|y - x\|_G^2. \end{array}$$

It follows that Condition ($GPCH$) holds with $b = (\lambda_{min}(G))^{-1}b' < 1$. $\square$

The next theorem concerns the case where $\Omega(x, y) = D(x)y$ for all $x, y \in H$, with $D(x)$ some positive definite matrix depending on $x$ not assumed to be symmetric. It is an extension of Corollary 2.10 of [99].

**Theorem 4.5** *Suppose that $F$ is $G$–differentiable and that there exists some positive constant $\Lambda$ such that $\|D(x)\| \leq \Lambda$ for all $x \in dom\ \varphi$. Suppose also that there exist a symmetric positive definite matrix $G$ and positive scalars $\nu$ and $\eta$ such that $\nu + \eta < \lambda_{min}(G)$, and for all $x, y \in dom\ \varphi$, $D(x) - G$ is positive semi-definite over dom $\varphi$ and*

$$\|\nabla F(x) - \nabla F(y)\| \leq \nu \quad \textit{and} \quad \|\nabla F(x) - D(x)\| \leq \eta.$$

*Then, if $x^0 \in dom\ \varphi$, $\{\varphi^k\}_{k \in I\!N}, \varphi \in \Gamma_0(H)$ are such that assumption (viii) of Proposition 4.2 holds, $F$ is weakly continuous on dom $\varphi$ and the functional $x \to \langle F(x), x\rangle$ is weakly lower semi–continuous on dom $\varphi$, the same conclusions as in Theorem 4.4 can be deduced.*

**Proof.** From the mean-value Theorem (see Proposition 1.6), we have:

$$\begin{aligned}&\|F(y) - F(x) - D(y)(y-x)\|\\ &\leq \sup_{0\leq t\leq 1} \|\nabla F(x + t(y-x)) - \nabla F(y)\|\, \|y - x\|\\ &+\|\nabla F(y) - D(y)\|\, \|y-x\|.\end{aligned}$$

Hence, it follows from the assumptions that

$$\|F(y) - F(x) - D(y)(y-x)\| \leq (\nu + \eta)\|y - x\|.$$

Then Proposition 4.6 ensures that Condition $(GPCH)$ is satisfied with $b = (\lambda_{min}(G))^{-1}(\nu + \eta)$. □

In the special case where $D(x) = \nabla F(x)$ for all $x$, this corollary could be interpreted as a global convergence result for the pertubed Newton method.

## 4.3 Local Convergence Results. Application to the Newton Method

As observed in the previous section, Theorem 4.5 could be applied to the Newton method by taking $D(x) = \nabla F(x)$ for all $x$. Nevertheless, this method is known to have a good convergence rate if considered in its local form. Motivated by this remark, we present in this section a local version of the general convergence theorem related to problems $(PAP^k)$, $k \in I\!N$. To obtain it, we suppose that assumptions $(vi)$ and $(vii)$ of Proposition 4.2 are satisfied only in a neighborhood of a solution of problem $(GVIP)$ and we impose an additional condition of closeness between the sequence $\{\varphi^k\}_{k\in I\!N}$ and the function $\varphi$. More precisely, our local convergence result can be stated as follows:

**Proposition 4.7** *Let $x^*$ be a solution of problem $(GVIP)$. Assume that the sequences $\{h^k\}_{k\in I\!N}$ and $\{\lambda_k\}_{k\in I\!N}$ satisfy assumptions (i)–(v) of Proposition 4.2 and that the following conditions are satisfied:*

**(a)** *there exist $\delta^* > 0$ and $\gamma > \bar{\lambda}/(2\beta')$ such that, for all $x, y$ in the closed ball $B(x^*, \delta^*)$ of center $x^*$ and radius $\delta^*$, the mapping $L(x, \cdot)$ is monotone on dom $\varphi$ and Lipschitz continuous on dom $\varphi$ uniformly in $x$ with Lipschitz constant $l > 0$ and,*

$$\text{if } \langle F(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0 \text{ holds, then}$$

$$\langle F(y) - L(y, y) + L(y, x), y - x\rangle + \varphi(y) - \varphi(x)$$

$$\geq \gamma \|(F(y) - L(y, y)) - (F(x) - L(y, x))\|^2;$$

**(b)** *$\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ are such that $\varphi^k \xrightarrow{M} \varphi$ and $\varphi \leq \varphi^k$ for all $k$. Moreover, there exist $\alpha > 1$ and a sequence $\{w^k\}_{k\in I\!N}$ satisfying (4.5) and such that the sequences $\{\alpha_k\}_{k\in I\!N}$ and $\{T^k\}_{k\in I\!N}$ defined by (4.16) and (4.17) satisfy the inequality $pT < \beta'\delta^{*2}/2$ where $T = \sum_{k=0}^{+\infty} T^k$ and $p = \lim_{k\to\infty} \prod_{j=0}^{k} \alpha_j^{-1}$.*

*Then, provided that $\Gamma^0(x^*, x^0) \leq \beta'\delta^{*2}/(2p) - T$, the sequence $\{x^k\}_{k\in I\!N}$ generated by problems $(PAP^k)$, $k \in I\!N$, stays in $B(x^*, \delta^*)$. In addition, if $F$ is weakly continuous on $B(x^*, \delta^*)$ and the functional $x \to \langle x, F(x)\rangle$ is weakly lower semi–continuous on $B(x^*, \delta^*)$, then any weak limit point of $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$ that belongs to $B(x^*, \delta^*)$. Moreover, if $x^*$ is the unique solution of $(GVIP)$ in $B(x^*, \delta^*)$, then $\{x^k\}_{k\in I\!N}$ weakly converges to $x^*$. In addition, if $F$ is strongly monotone on $B(x^*, \delta^*)$, then we also have that $\{x^k\}_{k\in I\!N}$ strongly converges to $x^*$.*

**Proof.** Let $\{\Gamma^k(x^*, \cdot)\}_{k\in I\!N}$ be the sequence of Lyapunov functions defined in (4.6). First, let us observe that if $x^k \in B(x^*, \delta^*)$ for each $k$, the same arguments as in the proofs of Propositions 4.2 and 4.3 allow us to conclude that any weak limit point of the sequence $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$.

In order to show that $x^k \in B(x^*, \delta^*)$ for each $k$, we consider the sequence $\{r_k\}_{k\in I\!N}$ defined by

$$r_0 = \delta \quad \text{and} \quad r_{k+1} = \alpha_k^{-1}(r_k + T^k), \quad k \in I\!N,$$

where $\delta = \beta'\delta^{*2}/(2p) - T$. It is easy to see that for all $k$, we have

$$\begin{aligned} r_{k+1} &= \delta \prod_{j=0}^{k} \alpha_j^{-1} + \sum_{j=0}^{k} (\prod_{t=j}^{k} \alpha_t^{-1}) T^j \\ &< \delta p + pT = (\beta'/2)\delta^{*2}. \end{aligned} \tag{4.25}$$

So, from (4.7) and (4.25), the iterate $x^k$ will belong to $B(x^*, \delta^*)$ as soon as $\Gamma^k(x^*, x^k) \le r_k$.
To end the proof, it remains to see that $\Gamma^k(x^*, x^k) \le r_k$ for all $k$. By induction, since this inequality is true for $k = 0$, we can suppose that it is true for $k$. Then $x^k \in B(x^*, \delta^*)$ and, since assumption $(a)$ holds on $B(x^*, \delta^*)$, we can conclude exactly as in the proof of Proposition 4.2 (see (4.8)) that

$$\Gamma^{k+1}(x^*, x^{k+1}) \le \alpha_k^{-1}(\Gamma^k(x^*, x^k) + T^k).$$

Then, by definition of $r_{k+1}$, we obtain that $\Gamma^{k+1}(x^*, x^{k+1}) \le r_{k+1}$. Since the last part of the theorem is straightforward, the proof is complete. $\square$

**Remark 4.8** When $\varphi$ is the indicator function of a nonempty closed convex subset $C$ of $H$, it follows from the assumptions on $\{h^k\}_{k\in I\!N}$ and $\{\lambda_k\}_{k\in I\!N}$ of Proposition 4.7, that for all $x \in C$,

$$\Gamma^0(x^*, x) \le (\Lambda'/2)\|x^* - x\|^2 + \overline{\lambda}\|F(x^*)\|\,\|x^* - x\|.$$

In this case, if we take $\epsilon > 0$ such that $(\Lambda'/2)\epsilon^2 + \overline{\lambda}\|F(x^*)\|\epsilon \le \beta'\delta^{*2}/(2p) - T$ and if we choose the starting point $x^0$ such that $x^0 \in C$ and $\|x^* - x^0\| \le \epsilon$, then $\Gamma^0(x^*, x^0) \le \beta'\delta^{*2}/(2p) - T$ and the conclusion of the preceding theorem holds.

As announced above, assumption $(b)$ imposes a condition of closeness between the sequence $\{\varphi^k\}_{k\in I\!N}$ and the function $\varphi$. This closeness is measured by the parameters $T$ and $p$. The more $T$ and $p$ are small, the more $\delta^*$ could be small. In order to prove the consistency of assumption $(b)$, it is important to see on examples that it is possible to build, for a $\delta^* > 0$ given, a sequence $\{\varphi^k\}_{k\in I\!N}$ which satisfies all the requirements of that assumption $(b)$. For that, we consider the barrier functions described in Example 3.3. We suppose that the functions $g_i, i = 1, \ldots, m$, and the subset $C$ are as in Example 3.3 with $g_0 = 0$. In addition, we impose that $C$ is bounded and we denote by $M$ a positive constant such that $-g_i(x) \leq M$ for all $x \in C$ and $1 \leq i \leq m$. For the sake of convenience, we slightly modify the barrier functions as follows:

$$\varphi^k(x) = -\nu_k^{-1} \sum_{i=1}^{m} \ln(\min(1/2, -g_i(x)/2M)), \quad \forall x \in I\!R^n, \quad \forall k \in I\!N, \tag{4.26}$$

where the sequence $\{\nu_k\}_{k\in I\!N}$ of barrier parameters is strictly increasing to $+\infty$, $\nu_k > 0$ for all $k$ and, by convention, $\ln(a) = -\infty$ when $a \leq 0$.
Suppose for a moment that the sequence $\{\nu_k\}_{k\in I\!N}$ has been chosen. Then, if we consider $\tilde{x} \in \text{int } C$ and define a sequence $\{w_k\}_{k\in I\!N}$ in $C$ by

$$w^k = x^* + \nu_k^{-1/2}(\tilde{x} - x^*), \ k \in I\!N,$$

we have successively, for all $k$,

$$\begin{aligned} w^k - x^* &= \nu_k^{-1/2}(\tilde{x} - x^*), && (4.27) \\ \varphi^k(w^k) &= -\nu_k^{-1} \sum_{i=1}^{m} \ln(\min(1/2, -g_i(w^k)/2M)) && \\ &\leq -\nu_k^{-1} \sum_{i=1}^{m} \ln(-g_i(\tilde{x})/(2M\nu_k^{1/2})) && \\ &= -\nu_k^{-1} \sum_{i=1}^{m} \ln(-g_i(\tilde{x})/2M) + m\nu_k^{-1} \ln(\nu_k^{1/2}), && (4.28) \end{aligned}$$

where the inequality follows from the convexity of $g_i, i = 1, \ldots, m$ and the fact that $x^* \in C$.
Gathering (4.17), (4.27) and (4.28), we obtain

$$T \leq c_1 \sum_{k=0}^{+\infty} \nu_k^{-1} + c_2 \sum_{k=0}^{+\infty} (k+1)^\alpha \nu_k^{-1} + c_3 \sum_{k=0}^{+\infty} \nu_k^{-1/2},$$

with

$$\begin{aligned} c_1 &= (1/2)(\Lambda'^2 \tau^{-1} + \overline{\lambda}\eta^{-1})\|\tilde{x} - x^*\|^2 - \overline{\lambda} \sum_{i=1}^{m} \ln(-g_i(\tilde{x})/2M), \\ c_2 &= l\overline{\lambda}\|\tilde{x} - x^*\|^2/2\theta, \\ c_3 &= \overline{\lambda}\|F(x^*)\|\|\hat{x} - x^*\| + m\overline{\lambda}. \end{aligned}$$

It is now clear that it is possible to choose the sequence $\{\nu_k\}_{k\in I\!N}$ in such a way that assumption $(b)$ be satisfied. $\square$

By the same process as before, we can deduce from Proposition 4.7 the local versions of the results of the preceding section. Since our aim is to provide a local version for methods of the Newton type, we will concentrate more particularly on local versions needed to achieve this goal. First, let us provide a local version to Theorem 4.3 for the case where $G = I$.

**Theorem 4.6** *Let $x^*$ be a solution of problem $(GVIP)$. Assume that the following conditions are satisfied:*

**(a)** *there exist $\delta^* > 0$ and $\gamma > 1/2$ such that, for all $x, y \in B(x^*, \delta^*)$, the mapping $\Omega(x,\cdot) - I$ is monotone on dom $\varphi$ and Lipschitz continuous on dom $\varphi$ uniformly in $x$ with Lipschitz constant $l > 0$ and,*

*if $\langle F(x), y - x\rangle + \varphi(y) - \varphi(x) \geq 0$ holds, then*

$$\langle F(y) - [\Omega(y,y) - y] + [\Omega(y,x) - x], y - x\rangle + \varphi(y) - \varphi(x)$$

$$\geq \gamma\|(F(y) - [\Omega(y,y) - y]) - (F(x) - [\Omega(y,x) - x])\|^2;$$

**(b)** *$\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ are such that $\varphi^k \xrightarrow{M} \varphi$ and $\varphi \leq \varphi^k$ for all $k$. Moreover, there exist $\alpha > 1$ and a sequence $\{w^k\}_{k\in I\!N}$ satisfying (4.5) and such that the inequality $pT < \delta^{*2}/2$ is satisfied with $T = \sum_{k=0}^{+\infty} T^k$, $p = \lim_{k\to\infty}\prod_{j=0}^{k}\alpha_j^{-1}$ and for all $k$,*

$$\begin{array}{rcl} \alpha_k & = & 1 - l\theta_k, \\ T^k & = & (1/2)[\tau^{-1} + \eta^{-1} + l\theta_k^{-1}]\|x^* - w^k\|^2 \\ & & +\|F(x^*)\|\|x^* - w^k\| + |\varphi^k(w^k) - \varphi(x^*)|, \end{array}$$

*where $\mu, \eta, \tau$ are such that $\mu + \eta - 2\gamma < 0$, $\tau + \mu^{-1} < 1$, $\theta^k = \theta(k+1)^{-\alpha}$ and $0 < \theta < l^{-1}$.*

*Then, provided that*

$$(1/2)\|x^* - x^0\|^2 + \langle F(x^*), x^0 - x^*\rangle + \varphi(x^0) - \varphi(x^*) \leq \delta^{*2}/(2p) - T,$$

*the same conclusions as in Proposition 4.7 are true.*

**Proof.** This result follows immediately from Proposition 4.7 with $h^k(y) = (1/2)y^T y$, $L(x^k, y) = \Omega(x^k, y) - y, \forall y, \forall k$ and $\lambda_k = 1, \forall k$, such that $\beta' = \Lambda' = 1$ and $\bar{\lambda} = 1$. $\square$

We can now state the local version of Theorem 4.4.

**Theorem 4.7** *Let $x^*$ be a solution of problem $(GVIP)$. Assume that the following conditions are satisfied:*

**(a)** *there exist a symmetric positive definite matrix $G$ and positive constants $\delta^*$ and $b < 1$ such that, for all $x, y \in B(x^*, \delta^*)$, $\Omega(x, \cdot) - G$ is monotone over dom $\varphi$ and Lipschitz continuous over dom $\varphi$ uniformly in $x$ with constant $l > 0$ and*

$$\|G^{-1}[F(y) - F(x) - (\Omega(y, y) - \Omega(y, x))]\|_G \le b\|y - x\|_G;$$

**(b)** *$\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ are such that $\varphi^k \xrightarrow{M} \varphi$ and $\varphi \le \varphi^k$ for all $k$. Moreover, there exist $\alpha > 1$ and a sequence $\{w^k\}_{k\in I\!N}$ satisfying (4.5) and such that the inequality $\widehat{p}\widehat{T} < \lambda_{min}(G)\delta^{*2}/2$ is satisfied with $\widehat{T} = \sum_{k=0}^{+\infty} \widehat{T^k}$, $\widehat{p} = \lim_{k\to\infty} \prod_{j=0}^{k} \widehat{\alpha_j}^{-1}$ and, for all $k \in I\!N$,*

$$\begin{aligned} \widehat{\alpha_k} &= 1 - l\theta_k \lambda_{min}(G)^{-1/2}, \\ \widehat{T^k} &= (1/2)(\tau^{-1} + \eta^{-1} + l\theta_k^{-1}\lambda_{min}(G)^{-1/2})\,\|x^* - w^k\|_G^2 \\ &\quad + \|G^{-1/2}F(x^*)\|\,\|x^* - w^k\|_G + |\varphi^k(w^k) - \varphi(x^*)|, \end{aligned}$$

*where $\mu, \eta, \tau$ are such that $\mu + \eta - (1 - b^2)/(1 + b)^2 < 1$, $\tau + \mu^{-1} < 1$, $\theta^k = \theta(k+1)^{-\alpha}$ and $0 < \theta < l^{-1}\lambda_{min}(G)^{1/2}$.*

*Then, provided that*

$$(1/2)\|x^* - x^0\|_G^2 + \langle F(x^*), x^0 - x^*\rangle + \varphi(x^0) - \varphi(x^*) \le \lambda_{min}(G)\,\delta^{*2}/(2\widehat{p}) - \widehat{T},$$

*the sequence $\{x^k\}_{k\in I\!N}$ generated by problems $(PAP^k)$, $k \in I\!N$, stays in $B(x^*, \delta^*)$. In addition, if $F$ is weakly continuous on $B(x^*, \delta^*)$ and the functional $x \to \langle F(x), x\rangle$ is weakly lower semi–continuous on $B(x^*, \delta^*)$, then the sequence $\{x^k\}_{k\in I\!N}$ strongly converges to $x^*$.*

**Proof.** As in the proof of Theorem 4.4, we consider the scaling $\widehat{x} = G^{1/2}x$ and we denote by $\widehat{F}, \widehat{\varphi}, \widehat{\varphi^k}, \widehat{\Omega}$ the scaled functions corresponding to $F, \varphi, \varphi^k$

and $\Omega$ respectively. Let us write conditions (a) and (b) for the variables $\widehat{x}$. Since

$$\begin{aligned}\|x - x^*\| &\leq (\lambda_{min}(G))^{-1/2}\|x - x^*\|_G \\ &= (\lambda_{min}(G))^{-1/2}\|\widehat{x} - \widehat{x^*}\|,\end{aligned}$$

we have that $x \in B(x^*, \delta^*)$ provided that $\widehat{x} \in B(\widehat{x^*}, \delta^*(\lambda_{min}(G))^{1/2})$. Moreover, if a function $J$ is Lipschitz continuous with Lipschitz constant $l$ in the $x$–space, then the function $\widehat{J}$ such that $\widehat{J}(\widehat{x}) = G^{-1/2}J(G^{-1/2}\widehat{x})$ is Lipschitz continuous with constant $l\,(\lambda_{min}(G))^{-1/2}$ in the $\widehat{x}$–space. Indeed,

$$\begin{aligned}\|\widehat{J}(\widehat{x}) - \widehat{J}(\widehat{y})\| &= \|J(x) - J(y)\|_{G^{-1}} \\ &\leq (\lambda_{max}(G^{-1}))^{1/2}\,\|J(x) - J(y)\| \\ &= (\lambda_{min}(G))^{-1/2}\|J(x) - J(y)\|.\end{aligned}$$

With these two observations, assumption (a) implies the following one in the $\widehat{x}$–space:

**($\widehat{\mathbf{a}}$)** there exist a symmetric positive definite matrix $G$ and positive constants $\delta^*$ and $b < 1$ such that, for all $\widehat{x}, \widehat{y} \in B(\widehat{x^*}, \delta^*(\lambda_{min}(G))^{1/2})$, $\widehat{\Omega}(\widehat{x}, \cdot) - I$ is monotone over dom $\varphi$ and Lipschitz continuous over dom $\varphi$ uniformly in $\widehat{x}$ with constant $l(\lambda_{min}(G))^{-1/2} > 0$ and

$$\|\widehat{F}(\widehat{y}) - \widehat{F}(\widehat{x}) - (\widehat{\Omega}(\widehat{y}, \widehat{y}) - \widehat{\Omega}(\widehat{y}, \widehat{x}))\| \leq b\|\widehat{y} - \widehat{x}\|.$$

In the sequel, we denote $\widehat{\delta^*} = \delta^*\,(\lambda_{min}(G))^{1/2}$ and $\widehat{l} = l\,(\lambda_{min}(G))^{-1/2}$. From the proof of Theorem 4.4, we know that condition ($\widehat{\text{a}}$) implies that $\widehat{F}$ is strongly monotone in $B(\widehat{x^*}, \widehat{\delta^*})$ and that the following condition holds:

**($\widehat{\widehat{\mathbf{a}}}$)** there exist positive constants $\widehat{\delta^*}$ and $\widehat{\gamma} = 1/2 + (1-b^2)/(2(1+b^2)) > 1/2$ such that, for all $\widehat{x}, \widehat{y} \in B(\widehat{x^*}, \widehat{\delta^*})$, $\widehat{\Omega}(\widehat{x}, \cdot) - I$ is monotone over dom $\varphi$ and Lipschitz continuous over dom $\varphi$ uniformly in $\widehat{x}$ with constant $\widehat{l}$ and

if $\langle \widehat{F}(\widehat{x}), \widehat{y} - \widehat{x}\rangle + \widehat{\varphi}(\widehat{y}) - \widehat{\varphi}(\widehat{x}) \geq 0$ holds, then

$$\langle \widehat{F}(\widehat{y}) - [\widehat{\Omega}(\widehat{y}, \widehat{y}) - \widehat{y}] + [\widehat{\Omega}(\widehat{y}, \widehat{x}) - \widehat{x}], \widehat{y} - \widehat{x}\rangle + \widehat{\varphi}(\widehat{y}) - \widehat{\varphi}(\widehat{x})$$

$$\geq \widehat{\gamma}\|(\widehat{F}(\widehat{y}) - [\widehat{\Omega}(\widehat{y}, \widehat{y}) - \widehat{y}]) - (\widehat{F}(\widehat{x}) - [\widehat{\Omega}(\widehat{y}, \widehat{x}) - \widehat{x}])\|^2;$$

On the other hand, condition (b) can be expressed in the following way:

**($\widehat{\mathrm{b}}$)** $\{\widehat{\varphi^k}\}_{k\in I\!N}, \widehat{\varphi} \in \Gamma_0(H)$ are such that $\widehat{\varphi^k} \overset{M}{\to} \widehat{\varphi}$ and $\widehat{\varphi} \leq \widehat{\varphi^k}$ for all $k$. Moreover, there exist $\alpha > 1$ and a sequence $\{\widehat{w^k}\}_{k\in I\!N} = \{G^{1/2}\, w^k\}_{k\in I\!N}$ satisfying

$$k^{\alpha}\|\widehat{w^k} - \widehat{x^*}\| \to 0 \text{ and } k^{\alpha}|\widehat{\varphi^k}(\widehat{w^k}) - \widehat{\varphi}(\widehat{x^*})| \to 0,$$

and such that the inequality $\widehat{p}\widehat{T} < \widehat{\delta^{*2}}/2$ is satisfied with $\widehat{T} = \sum_{k=0}^{+\infty} \widehat{T^k}$, $\widehat{p} = \lim_{k\to\infty} \prod_{j=0}^{k} \widehat{\alpha_j}^{-1}$ and, for all $k \in I\!N$,

$$\begin{array}{rcl} \widehat{\alpha_k} & = & 1 - \widehat{l}\theta_k, \\ \widehat{T^k} & = & (1/2)(\tau^{-1} + \eta^{-1} + \widehat{l}\theta_k^{-1})\|\widehat{x^*} - \widehat{w^k}\|^2 \\ & & +\|\widehat{F}(\widehat{x^*})\|\,\|\widehat{x^*} - \widehat{w^k}\| + |\widehat{\varphi^k}(\widehat{w^k}) - \widehat{\varphi}(\widehat{x^*})|, \end{array}$$

where $\mu, \eta, \tau$ are such that $\mu + \eta - 2\widehat{\gamma} < 0,\ \tau + \mu^{-1} < 1$, $\theta^k = \theta(1+k)^{-\alpha}$ and $0 < \theta < \widehat{l}^{-1}$.

Finally, it can be easily seen that the inequality

$$(1/2)\|x^* - x^0\|_G^2 + \langle F(x^*), x^0 - x^*\rangle + \varphi(x^0) - \varphi(x^*) \leq \lambda_{min}(G)\,\delta^{*2}/(2\widehat{p}) - \widehat{T}$$

can be expressed as

$$(1/2)\|\widehat{x^*} - \widehat{x^0}\|^2 + \langle \widehat{F}(\widehat{x^*}), \widehat{x^0} - \widehat{x^*}\rangle + \widehat{\varphi}(\widehat{x^0}) - \widehat{\varphi}(\widehat{x^*}) \leq \widehat{\delta^{*2}}/(2\widehat{p}) - \widehat{T}.$$

This last remark together with conditions ($\widehat{\mathrm{a}}$) and ($\widehat{\mathrm{b}}$) allow us to apply Theorem 4.6 in the $\widehat{x}$–space knowing that $\widehat{x}^*$ is the unique solution of $(\widehat{GVIP})$ in $B(\widehat{x^*}, \widehat{\delta^*})$. After translating these conclusions in the $x$–space, the proof is then complete. $\square$

When $\Omega(x, y) = D(x)y$ for all $x, y \in H$ with $D(x)$ some positive definite matrix, the local convergence result corresponding to Theorem 4.5 can be stated in the following way:

**Theorem 4.8** *Suppose that $F$ is $G$–differentiable and let $x^*$ be a solution of problem $(GVIP)$. Assume that the following conditions are satisfied:*

**(a)** *there exist a symmetric positive definite matrix $G$ and positive scalars $\delta^*, \nu$ and $\eta$ with $\nu + \eta < \lambda_{min}(G)$ such that for all $x, y \in B(x^*, \delta^*)$, $\|D(x)\| \leq \Lambda$ for some $\Lambda > 0$, the matrix $D(x) - G$ is positive semi-definite over dom $\varphi$ and*

$$\|\nabla F(x) - \nabla F(y)\| \leq \nu \quad \textit{and} \quad \|\nabla F(x) - D(x)\| \leq \eta;$$

**(b)** *assumption (b) of Theorem 4.7 holds with* $l = \Lambda + \lambda_{max}(G)$.

*Then, the same conclusions as in Theorem 4.7 can be deduced.*

**Proof.** This result follows directly from Theorem 4.7 by using the same arguments as in the proof of Theorem 4.5. □

To end this section, we can consider more specially the application of this corollary to the Newton method. Suppose that $F$ is continuously G–differentiable. Then, the sequence generated by the perturbed Newton method is defined by the following problems:

$$(NM^k) \begin{cases} \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle F(x^k) + \nabla F(x^k)(x^{k+1} - x^k), x - x^{k+1} \rangle + \varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

The corresponding local convergence result can be expressed as follows:

**Corollary 4.1** *Let* $x^*$ *be a solution of problem* $(GVIP)$. *Assume that the following conditions are satisfied:*

**(a)** *there exist a symmetric positive definite matrix* $G$ *and positive scalars* $\delta^*$ *and* $\nu < \lambda_{min}(G)$ *such that, for all* $x, y \in B(x^*, \delta^*)$, *the matrix* $\nabla F(x) - G$ *is positive semi-definite on dom* $\varphi$ *and* $\|\nabla F(x) - \nabla F(y)\| \leq \nu$;

**(b)** *assumption (b) of Theorem 4.7 holds with* $l = \|\nabla F(x^*)\| + \nu + \lambda_{max}(G)$.

*Then, the same conclusions as in Theorem 4.7 are obtained.*

**Proof.** This is an obvious consequence of Theorem 4.8 where we take $D(x) = \nabla F(x)$. □

Finally, observe that, when $\varphi^k = \varphi$ for all $k$, the parameters $T^k$ and $\widehat{T^k}$ are equal to zero for all $k$. Hence, in that case, $T = \widehat{T} = 0$ and assumption $(b)$ of Proposition 4.7, Theorems 4.6–4.8 and Corollary 4.1 is obviously satisfied. This allows us to recover well-known local convergence results as the following one due to Pang and Chan ([99], Corollary 2.6):

**Corollary 4.2** *($H = I\!R^n$) If* $x^*$ *is a solution of problem* $(GVIP)$ *and* $\nabla F(x^*)$ *is a positive definite matrix, then, provided that* $x^0$ *is chosen close enough to* $x^*$, *the sequence generated by problems* $(NM^k)$, $k \in I\!N$, *with* $\varphi^k = \varphi$ *for all* $k$ *strongly converges to* $x^*$.

**Proof.** Since $\nabla F(x^*)$ is a positive definite matrix, there exists a positive constant $\mu$ such that, for all $s \in H$:

$$s^T \nabla F(x^*) s \geq \mu \|s\|^2.$$

If we take $\mu'$ such that $0 < \mu' < \mu$ and $G = \mu' I$, then $\nabla F(x^*) - G$ is positive definite and by the continuity of $\nabla F$, there exists a neighborhood $N_1$ of $x^*$ such that $\nabla F(x) - G$ is positive semi-definite for all $x$ in $N_1$.
Next, using again the continuity of $\nabla F$, for any positive constant $\nu < \mu'$, there exists a neighborhood $N_2$ of $x^*$ such that $\|\nabla F(x) - \nabla F(y)\| \leq \nu$ for all $x, y$ in $N_2$. Hence assumption $(a)$ of Corollary 4.1 is satisfied for a ball $B(x^*, \delta^*)$ contained in $N_1 \cap N_2$. Since assumption $(b)$ of Corollary 4.1 holds obviously, the conclusion follows immediately. □

# Chapter 5

# Convergence of the Perturbed Auxiliary Problem Method for Multivalued Mappings

In this chapter, the operator $F$ is multivalued and the auxiliary operator is fixed and symmetric. So, at iteration $k$, the vector $x^{k+1}$ is the solution of the following perturbed symmetric auxiliary subproblem:

$$(PSAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\nabla K(x^{k+1}) - \nabla K(x^k)), x - x^{k+1}\rangle \\ \qquad\qquad\qquad +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

In a first part, we provide convergence results for this perturbed scheme under the same kind of assumptions on $F$ as in [134] (see Theorem 2.6). In a second part, we relax this scheme by allowing an inexact computation of an element of $F(x^k)$. This is made by taking $r(x^k)$ in an enlargement of $F$ at $x^k$. We discuss the choice of an adequate enlargement and the corresponding convergence conditions. In the particular case of nondifferentiable convex optimization, the $\epsilon$–subdifferential can be chosen as the enlargement of the subdifferential operator so that our scheme is a generalization of the

projected inexact subgradient procedure analyzed in [4].
Most results of the first section of this chapter are presented in [113]. The results of the third section are a generalization of those appeared in [111] which deals with the more restrictive case where $F$ is supposed to be strongly monotone.

## 5.1 Convergence Results

In this section, we present successively the assumptions ensuring that the sequence $\{x^k\}_{k\in I\!N}$ generated by solving problems $(PSAP^k)$ is bounded, admits at least one weak limit point which is a solution of problem $(GVIP)$, weakly and then strongly converges to a solution of problem $(GVIP)$.

In the multivalued case, the sequence $\{\lambda_k\}_{k\in I\!N}$ is chosen as, for example, in [4], i.e.

$$\begin{cases} \lambda_k = \mu_k/\eta_k, \forall k \in I\!N, \text{ with } \{\mu_k\}_{k\in I\!N} \text{ a sequence of positive numbers,} \\ \text{and } \eta_k = \begin{cases} \max\{1, \|r(x^0)\|\}, \text{ if } k = 0, \\ \max\{\eta_{k-1}, \|r(x^k)\|\}, \text{ if } k \geq 1. \end{cases} \end{cases}$$

The introduction of the sequence $\{\eta_k\}_{k\in I\!N}$ allows us to prove that the sequence $\{x^k\}_{k\in I\!N}$ is bounded without any additional assumption on the mapping $F$. The case $\eta_k = 1$ for all $k$ is treated in Remark 5.3 at the end of this section. Moreover, as it is classically assumed in the multivalued case, the positive sequence $\{\mu_k\}_{k\in I\!N}$ will be such that

$$\sum_{k=0}^{+\infty} \mu_k^2 < +\infty \quad \text{and} \quad \sum_{k=0}^{+\infty} \mu_k = +\infty.$$

The sequence of Lyapunov functions $\{\Gamma^k(x^*, \cdot)\}_{k\in I\!N}$ considered here is defined on H by

$$\begin{aligned} \Gamma^k(x^*, x) &= K(x^*) - K(x) - \langle \nabla K(x), x^* - x\rangle \\ &\quad + (\mu_k/\eta_k)[\langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*)], \end{aligned} \tag{5.1}$$

where $x^*$ denotes a solution of problem $(GVIP)$ and $r(x^*)$ is an element in $F(x^*)$ such that $\langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*) \geq 0$, for all $x$ in $H$. The next lemma gives an upper bound on $\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k)$.

**Lemma 5.1** *Assume that $F$ is a monotone multivalued mapping defined on $H$, that problem $(GVIP)$ admits at least one solution denoted by $x^*$, and that the following conditions are satisfied:*

**(i)** $K : H \to I\!R$ *is continuously differentiable and strongly convex with modulus $\beta > 0$ over dom $\varphi$;*

**(ii)** $\nabla K$ *is a Lipschitz continuous mapping with Lipschitz constant $\Lambda$ over dom $\varphi$;*

**(iii)** $\{\mu_k\}_{k\in I\!N}$ *is a nonincreasing sequence of positive numbers;*

**(iv)** $\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ *are such that $\varphi \leq \varphi^k$ for all $k$, and there exists a sequence $\{w^k\}_{k\in I\!N}$ such that*

$$\sum_{k=0}^{+\infty} \|w^k - x^*\| < +\infty \quad \text{and} \quad \sum_{k=0}^{+\infty} |\varphi^k(w^k) - \varphi(x^*)| < +\infty. \tag{5.2}$$

*Then, if $\{x^k\}_{k\in I\!N}$ denotes the sequence generated by solving subproblems $(PSAP^k)$, we have for all $k \in I\!N$,*

$$\begin{aligned} \Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq -c\|x^{k+1} - x^k\|^2 + T^k + \mu_k^2 u \\ -(\mu_k/\eta_k)[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)], \end{aligned} \tag{5.3}$$

*with $c, u > 0$, $T^k \geq 0$, and $\sum_{k=0}^{+\infty} T^k < +\infty$.*

**Proof.** The necessary and sufficient optimality conditions satisfied by $x^*$ and $x^{k+1}$ are respectively

$$\langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*) \geq 0, \ \forall x \in H, \tag{5.4}$$

and

$$\begin{aligned} &\langle \eta_k^{-1} r(x^k) + \mu_k^{-1}(\nabla K(x^{k+1}) - \nabla K(x^k)), x - x^{k+1}\rangle \\ &+\eta_k^{-1}(\varphi^k(x) - \varphi^k(x^{k+1})) \geq 0, \forall x \in H, \ \text{with } r(x^k) \in F(x^k). \end{aligned} \tag{5.5}$$

We consider the sequence of Lyapunov functions $\{\Gamma^k(x^*, \cdot)\}_{k\in I\!N}$ defined in (5.1). From the strong convexity of $K$ and inequality (5.4), we obtain that, for all $x \in$ dom $\varphi$, and all $k \in I\!N$,

$$\Gamma^k(x^*, x) \geq (\beta/2)\|x - x^*\|^2. \tag{5.6}$$

Using the definition of the Lyapunov function and the facts that $\mu_{k+1} \le \mu_k$, and $\eta_{k+1} \ge \eta_k$ for all $k \in I\!N$, we can write

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \le \Gamma^k(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) = s_1 + s_2 + s_3, \quad (5.7)$$

$$\begin{array}{lrcl} \text{with} & s_1 & = & K(x^k) - K(x^{k+1}) + \langle \nabla K(x^k), x^{k+1} - x^k \rangle, \\ & s_2 & = & \langle \nabla K(x^k) - \nabla K(x^{k+1}), x^* - x^{k+1} \rangle, \\ & s_3 & = & (\mu_k/\eta_k)[\langle r(x^*), x^{k+1} - x^k \rangle + \varphi(x^{k+1}) - \varphi(x^k)]. \end{array}$$

For $s_1$, we derive easily from the strong convexity of $K$ that

$$s_1 \le -(\beta/2)\|x^{k+1} - x^k\|^2. \quad (5.8)$$

Now, using the sequence $\{w^k\}_{k \in I\!N}$ given in assumption (iv), we can write $s_2$ as the sum of the two following terms:

$$\begin{array}{rcl} s_{21} & = & \langle \nabla K(x^k) - \nabla K(x^{k+1}), x^* - w^k \rangle, \\ s_{22} & = & \langle \nabla K(x^k) - \nabla K(x^{k+1}), w^k - x^{k+1} \rangle. \end{array}$$

From the Lipschitz continuity of $\nabla K$, we deduce that

$$s_{21} \le \Lambda \|x^{k+1} - x^k\| \, \|x^* - w^k\| \le (\tau/2)\|x^{k+1} - x^k\|^2 + (\Lambda^2/(2\tau))\|x^* - w^k\|^2, \quad (5.9)$$

where the second inequality holds for any $\tau > 0$.
Using (5.5) with $x = w^k$, we obtain

$$\begin{array}{rcl} s_{22} & \le & (\mu_k/\eta_k)[\langle r(x^k), w^k - x^{k+1} \rangle + \varphi^k(w^k) - \varphi^k(x^{k+1})] \\ & = & (\mu_k/\eta_k)[\langle r(x^k), w^k - x^* \rangle \\ & & + \langle r(x^k), x^* - x^k \rangle + \varphi(x^*) - \varphi(x^k) \\ & & + \langle r(x^k), x^k - x^{k+1} \rangle \\ & & + \varphi^k(w^k) - \varphi(x^*) + \varphi(x^k) - \varphi^k(x^{k+1})], \end{array}$$

so that

$$\begin{array}{rcl} s_{22} + s_3 & = & (\mu_k/\eta_k)[\langle r(x^k), w^k - x^* \rangle \\ & & + \langle r(x^k), x^* - x^k \rangle + \varphi(x^*) - \varphi(x^k) \end{array}$$

$$+\langle r(x^k), x^k - x^{k+1}\rangle$$

$$+\langle r(x^*), x^{k+1} - x^k\rangle$$

$$+\varphi^k(w^k) - \varphi(x^*) + \varphi(x^{k+1}) - \varphi^k(x^{k+1})]. \qquad (5.10)$$

From the definition of the sequence $\{\eta_k\}_{k\in I\!N}$ and the fact that $\mu_k \leq \mu_0$ for all $k$, we have

$$(\mu_k/\eta_k)\langle r(x^k), w^k - x^*\rangle \leq \mu_0\|w^k - x^*\|, \qquad (5.11)$$

$$\begin{aligned}(\mu_k/\eta_k)\langle r(x^k), x^k - x^{k+1}\rangle &\leq \mu_k\|x^{k+1} - x^k\| \\ &\leq \mu_k^2/(2\gamma) + (\gamma/2)\|x^{k+1} - x^k\|^2, \qquad (5.12)\end{aligned}$$

$$\begin{aligned}(\mu_k/\eta_k)\langle r(x^*), x^{k+1} - x^k\rangle &\leq \mu_k\|r(x^*)\|\ \|x^{k+1} - x^k\| \\ &\leq (\mu_k^2/2\mu)\|r(x^*)\|^2 \\ &\quad +(\mu/2)\|x^{k+1} - x^k\|^2, \qquad (5.13)\end{aligned}$$

with $\gamma$ and $\mu$ any positive numbers.
Gathering the fact that $\varphi \leq \varphi^k$ for all $k$ with inequalities (5.7)–(5.13) and rearranging the terms, we deduce that inequality (5.3) holds with

$$\begin{array}{lcl} c &=& (1/2)(\beta - \tau - \gamma - \mu), \\ T^k &=& \mu_0\|w^k - x^*\| + \mu_0|\varphi^k(w^k) - \varphi(x^*)| + (\Lambda^2/(2\tau))\|w^k - x^*\|^2, \\ u &=& (1/(2\gamma)) + (1/(2\mu))\|r(x^*)\|^2, \\ \tau, \quad \gamma, & & \mu > 0 \ \text{ such that } \tau + \gamma + \mu < \beta. \end{array}$$

Since the sequence $\{w^k\}_{k\in I\!N}$ has been chosen such that (5.2) holds, we have that $\sum_{k=0}^{+\infty} T^k < +\infty$ and the proof is complete. □

The next theorem gives conditions that ensure boundedness of the sequence $\{x^k\}_{k\in I\!N}$.

**Theorem 5.1** *Assume that all assumptions of Lemma 5.1 hold.*
*If* $\sum_{k=0}^{+\infty} \mu_k^2 < +\infty$, *then provided that* $x^0 \in \mathrm{dom}\ \varphi$, *the sequence* $\{x^k\}_{k\in I\!N}$ *is bounded. Moreover,*

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty, \ \textit{and}$$

$$\sum_{k=0}^{+\infty}(\mu_k/\eta_k)[\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*)] < +\infty.$$

**Proof.** Since $r(x^k) \in F(x^k)$ for all $k$, $x^*$ is a solution of problem $(GVIP)$ and $F$ is monotone, we have that

$$(\mu_k/\eta_k)[\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*)] \geq 0. \tag{5.14}$$

So, we derive from (5.3) that

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq T^k + u\,\mu_k^2. \tag{5.15}$$

Since $\sum_{k=0}^{+\infty} T^k$, $\sum_{k=0}^{+\infty} \mu_k^2$ are convergent series, it follows that $\{\Gamma^k(x^*, x^k)\}_{k\in\mathbb{N}}$ is a Cauchy sequence. This implies in turn that it is convergent in $H$. Using inequality (5.6), we can conclude that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded. Then, rearranging the terms of inequality (5.3) as follows

$$c\|x^{k+1} - x^k\|^2 + (\mu_k/\eta_k)[\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*)]$$

$$\leq \Gamma^k(x^*, x^k) - \Gamma^{k+1}(x^*, x^{k+1}) + T^k + \mu_k^2\, u,$$

we obtain that $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$ and

$$\sum_{k=0}^{+\infty}(\mu_k/\eta_k)[\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*)] < +\infty.$$

□

To prove that at least one weak limit point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a solution of problem $(GVIP)$, we have to impose that $F$ is paramonotone, bounded on bounded subsets of dom $\varphi$ and weakly closed on dom $\varphi$. If $F$ is bounded on bounded subsets and weakly closed, we can prove that at least one weak limit point $\bar{x}$ of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by the algorithm satisfies $\langle \bar{r}, x^* - \bar{x} \rangle + \varphi(x^*) - \varphi(\bar{x}) \geq 0$, with $\bar{r} \in F(\bar{x})$ and $x^*$ a solution of problem $(GVIP)$. Then paramonotonicity is used to deduce that $\bar{x}$ is a solution of problem $(GVIP)$ (see Proposition 1.26). Observe that, when $F$ is monotone, we have that $F$ is locally bounded in the interior of its domain (see Proposition 1.10) so that it is bounded on bounded subsets in finite dimension. On the other hand, if $F$ is maximal monotone, then it is weakly–strongly (or strongly–weakly) closed (see Proposition 1.14). In finite dimension, this implies that it is (strongly) closed.

We obtain the following convergence result.

**Theorem 5.2** *Suppose that all assumptions of Lemma 5.1 are satisfied and that the sequence $\{x^k\}_{k\in I\!N}$ is bounded. If $\lim_{k\to\infty}\mu_k = 0$, $\sum_{k=0}^{+\infty}\mu_k = +\infty$, $F$ is paramonotone over dom $\varphi$, bounded on bounded subsets of dom $\varphi$ and weakly closed on dom $\varphi$, then at least one weak limit point of the sequence $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$.*

**Proof.** Since the sequence $\{x^k\}_{k\in I\!N} \subset$ dom $\varphi$ is bounded and $F$ is bounded on bounded subsets of dom $\varphi$, we derive that the sequence $\{r(x^k)\}_{k\in I\!N}$ is also bounded and there exists a constant $\bar{\eta} > 1$ such that $\|r(x^k)\| \leq \bar{\eta}$ for all $k$. Therefore, we have also that $\eta_k \leq \bar{\eta}$ for all $k$.
Now, we show that

$$\underline{\lim}_{k\to\infty}[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] = 0.$$

Since $r(x^k) \in F(x^k)$ for all $k$ and $F$ is monotone, we have that

$$\underline{\lim}_{k\to\infty}[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] \geq 0.$$

To prove the equality, assume for contradiction that the inferior limit is strictly positive i.e., there exist $\delta > 0$ and $k_0 \in I\!N$ such that

$$\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*) > \delta,\ \forall k \geq k_0.$$

Since the sequence $\{\mu_k\}_{k\in I\!N}$ converges to zero, there also exists $k_1 \in I\!N$ such that

$$\mu_k u < \delta/(2\bar{\eta}),\ \forall k \geq k_1.$$

So, we derive from these two last inequalities and from (5.3) that

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq T^k - \mu_k\delta/(2\bar{\eta}),\ \forall k \geq \bar{k} = \max(k_0, k_1).$$

Summing up, this gives for all $N \geq \bar{k}$:

$$0 \leq \Gamma^{N+1}(x^*, x^{N+1}) \leq \Gamma^{\bar{k}}(x^*, x^{\bar{k}}) + \sum_{k=\bar{k}}^{N} T^k - (\delta/(2\bar{\eta}))\sum_{k=\bar{k}}^{N}\mu_k.$$

If we take the limit on $N$, we obtain a contradiction because the series $\sum_{k=\bar{k}}^{+\infty} T^k$ is convergent and the series $\sum_{k=0}^{+\infty}\mu_k$ is divergent.
Hence, we deduce that there exists a subsequence $\{x^{k'}\}_{k'\in K\subset I\!N}$ such that

$$\lim_{k'\to\infty}\langle r(x^{k'}), x^{k'} - x^*\rangle + \varphi(x^{k'}) - \varphi(x^*) = 0.$$

In addition, since $F$ is monotone and weakly closed, and $\varphi$ is weakly lower semi–continuous (see Proposition 1.1), there exist subsequences

$$\{x^{k''}\}_{k''\in K'\subset K}, \text{ and } \{r(x^{k''})\}_{k''\in K'\subset K} \text{ such that}$$

$$x^{k''} \rightharpoonup \bar{x},\ r(x^{k''}) \rightharpoonup \bar{r},\ \bar{r} \in F(\bar{x}),\ \text{and}$$

$$\underline{\lim}_{k''\to\infty}[\langle r(x^{k''}), x^{k''} - \bar{x}\rangle + \varphi(x^{k''}) - \varphi(\bar{x})] \geq 0.$$

This implies that

$$0 = \lim_{k''\to\infty}[\langle r(x^{k''}), x^{k''} - x^*\rangle + \varphi(x^{k''}) - \varphi(x^*)] \geq \langle \bar{r}, \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*).$$

Since $F$ is paramonotone over dom $\varphi$, it follows from Proposition 1.26 that $\bar{x}$ solves problem $(GVIP)$. □

This result provides informations on the limit behavior of the generated sequence under very weak restrictions on $F$. Now, to obtain that each weak limit point of $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$, we will use the concept of gap function (see, for example, [11]). We recall that a function $l$ : dom $\varphi \longrightarrow I\!R \cup \{+\infty\}$ is a *gap function* with respect to problem $(GVIP)$ if

$l(x) \geq 0$ for all $x \in$ dom $\varphi$ and

$l(\bar{x}) = 0$ if and only if $\bar{x}$ is a solution of problem $(GVIP)$.

In our context, the usefulness of a gap function appears in the next proposition. In that proposition and in the sequel, we denote by $E$ a set satisfying

$$\{x^k\}_{k\in I\!N} \subseteq E \subseteq \text{dom } \varphi,$$

where $\{x^k\}_{k\in I\!N}$ is the sequence generated by solving subproblems $(PSAP^k)$.

**Proposition 5.1** *Let $l$ be a gap function with respect to problem $(GVIP)$. If $l$ is weakly lower semi–continuous on $E$ and if $l(x^k) \to 0$, then any weak limit point of the sequence $\{x^k\}_{k\in I\!N}$ generated by the algorithm is a solution of problem $(GVIP)$.*

**Proof.** First, notice that the sequence $\{x^k\}_{k\in I\!N}$ is contained in $E$. Let $\bar{x}$ be a weak limit point of this sequence such that the subsequence $\{x^{k'}\}_{k'\in K\subset I\!N}$ weakly converges to $\bar{x}$. Then,

$$0 = \lim_{k\to+\infty} l(x^k) = \underline{\lim}_{k'\to+\infty} l(x^{k'}) \geq l(\bar{x}) \geq 0,$$

i.e., $l(\bar{x}) = 0$ and $\bar{x}$ is a solution of problem $(GVIP)$. $\square$

To prove that $l(x^k) \to 0$, we will use the following lemma due to Cohen and Zhu.

**Lemma 5.2** (See [34], Lemma 4) *If $l$ is a Lipschitz continuous function on $\{x^k\}_{k\in I\!N}$ and if $\{\mu_k\}_{k\in I\!N}$ is a sequence of positive numbers satisfying*
*(a)* $\sum_{k=0}^{+\infty} \mu_k = +\infty$;
*(b)* $\sum_{k=0}^{+\infty} \mu_k\, l(x^k) < +\infty$;
*(c)* $\exists\, \delta > 0$ *such that* $\forall k \in I\!N,\ \|x^{k+1} - x^k\| \le \delta\mu_k$,
*then* $l(x^k) \to 0$.

Before enouncing the main convergence results, we give three existence results for gap functions weakly lower semi–continuous on $E$ and Lipschitz continuous on bounded subsets of $E$ and then we present conditions under which assumptions $(b)$ and $(c)$ of Lemma 5.2 are satisfied.

**Proposition 5.2** *Let $x^*$ denote any solution of problem $(GVIP)$.*
*(a) If $F$ is paramonotone on dom $\varphi$ and $F(x)$ is a bounded and weakly closed subset of $H$ for all $x \in$ dom $\varphi$, then*

$$\mathbf{l(x)} = \inf_{\mathbf{r(x)\in F(x)}} \langle \mathbf{r(x), x - x^*}\rangle + \varphi(\mathbf{x}) - \varphi(\mathbf{x^*})$$

*is a gap function.*
*(b) If, in addition, $F$ and $\varphi$ are Lipschitz continuous on bounded subsets of $E$, then $l$ is Lipschitz continuous on bounded subsets of $E$.*
*(c) If, in addition, $F$ is weakly closed on dom $\varphi$, then $l$ is weakly lower semi–continuous on $E$.*

**Proof.** (a) Since $F$ is monotone and $x^*$ is a solution of problem $(GVIP)$, for each $x \in$ dom $\varphi$ and any $r(x) \in F(x)$, we have

$$\begin{aligned}&\langle r(x), x - x^*\rangle + \varphi(x) - \varphi(x^*)\\ &= \langle r(x) - r(x^*), x - x^*\rangle + \langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*) \ge 0.\end{aligned}$$

So, using the definition of $l$, we obtain that $l(x) \ge 0$. Now if $\bar{x}$ is a solution of $(GVIP)$, then we have immediately that

$$l(\bar{x}) \le \langle r(\bar{x}), \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) \le 0 \le l(\bar{x}).$$

So, $l(\bar{x}) = 0$. Conversely, suppose that $l(\bar{x}) = 0$. Then, by definition of the infimum, there exists a sequence $\{r_k\}_{k\in I\!N}$ contained in $F(\bar{x})$ such that, for all $k \geq 1$,

$$0 \leq \langle r_k, \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) < 1/k.$$

Since the subset $F(\bar{x})$ is bounded and weakly closed, there exists a subsequence of $\{r_k\}_{k\in I\!N}$ that weakly converges to some $r \in F(\bar{x})$. Then $0 \leq \langle r, \bar{x} - x^*\rangle + \varphi(\bar{x}) - \varphi(x^*) \leq 0$, and by Proposition 1.26, $\bar{x}$ is a solution of $(GVIP)$ because $F$ is paramonotone.

(b) Let $B$ be a bounded subset of $E$ and $c_1 > 0$ be such that $\|x\| \leq c_1$ for all $x \in B$. Since $\varphi$ is Lipschitz continuous on $B$, it is sufficient to prove that there exists $L_1 > 0$ such that, for all $x, y \in B$,

$$\inf_{r\in F(x)} \langle r, x - x^*\rangle + \sup_{s\in F(y)} \langle s, x^* - y\rangle \leq L_1\|x - y\|. \tag{5.16}$$

Let $x, y \in B, \epsilon > 0$ and $s \in F(y)$. Since $F$ is Lipschitz continuous on $B$, there exists $L > 0$ such that $e(F(y), F(x)) \leq L\|x - y\|$, we have

$$\inf_{r\in F(x)} \|r - s\| \leq L\|x - y\|.$$

So, there exists $r \in F(x)$ such that $\|r-s\| \leq L\|x-y\|+\epsilon/(c_1+\|x^*\|)$. Then

$$\begin{aligned}\langle r, x - x^*\rangle + \langle s, x^* - y\rangle &= \langle r, x - y\rangle + \langle r - s, y - x^*\rangle \\ &\leq \|r\|\|x - y\| + \|r - s\|\|y - x^*\| \\ &\leq \|r\|\|x - y\| + L\|x - y\|(c_1 + \|x^*\|) + \epsilon.\end{aligned} \tag{5.17}$$

Moreover, by Lemma 1.1, $F$ is bounded on $B$ and consequently, there exists $c > 0$ such that $\|r\| \leq c$ for all $r \in F(x)$ and $x \in B$. Then, from (5.17), we deduce that

$$\inf_{r\in F(x)} \langle r, x - x^*\rangle + \langle s, x^* - y\rangle \leq L_1\|x - y\| + \epsilon,$$

where $L_1 = c+L(c_1+\|x^*\|)$. Since this inequality is satisfied for all $s \in F(y)$ and $\epsilon > 0$, we obtain (5.16).

(c) Suppose that $F$ is weakly closed on dom $\varphi$. Since $\varphi$ is weakly lower semi–continuous on dom $\varphi$, we have only to prove that

$$l^1(x) \equiv \inf_{r(x)\in F(x)} \langle r(x), x - x^*\rangle$$

is weakly lower semi–continuous on $E$. Let $x^k \rightharpoonup \bar{x}$ with $x^k \in E$, and let $\bar{l}^1$ be a limit point of the sequence $\{l^1(x^k)\}_{k\in I\!N}$. We have to prove that $\bar{l}^1 \geq l^1(\bar{x})$. Without loss of generality, we can suppose that $l^1(x^k) \to \bar{l}^1$. Let $\epsilon > 0$, by definition of the infimum, for each $k$, there exists $r(x^k) \in F(x^k)$ such that

$$\langle r(x^k), x^k - x^* \rangle \leq l^1(x^k) + \epsilon. \tag{5.18}$$

Since the sequence $\{x^k\}_{k\in I\!N}$ is bounded and contained in $E$, and since $F$ is bounded on bounded subsets of $E$ (see Lemma 1.1), the sequence $\{r(x^k)\}_{k\in I\!N}$ is bounded and thus there exists a subsequence $\{r(x^{k'})\}_{k'\in K}$ weakly converging to some $\bar{r}$. Since $F$ is weakly closed, it follows that $\bar{r} \in F(\bar{x})$. Now, $F$ being monotone, we have that $\langle r(x^{k'}) - r(\bar{x}), x^{k'} - \bar{x}\rangle \geq 0$ and thus that

$$\langle r(x^{k'}), x^{k'} - x^* \rangle \geq \langle r(\bar{x}), x^{k'} - \bar{x} \rangle + \langle r(x^{k'}), \bar{x} - x^* \rangle. \tag{5.19}$$

Gathering (5.18) and (5.19), we obtain

$$l^1(x^{k'}) + \epsilon \geq \langle r(\bar{x}), x^{k'} - \bar{x} \rangle + \langle r(x^{k'}), \bar{x} - x^* \rangle. \tag{5.20}$$

Passing to the limit in (5.20) and noticing that $\langle \bar{r}, \bar{x} - x^* \rangle \geq l^1(\bar{x})$, we have that $\bar{l}^1 + \epsilon \geq l^1(\bar{x})$. Since $\epsilon$ is arbitrary, we have that $\bar{l}^1 \geq l^1(\bar{x})$ and consequently $l$ is weakly lower semi–continuous on $E$. □

**Proposition 5.3** *Let $x^*$ denote any solution of problem $(GVIP)$. If $F = \partial f, f \in \Gamma_0(H)$ and dom $\varphi \subseteq$ int(dom$f$), then*

$$\mathbf{l}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \varphi(\mathbf{x}) - \mathbf{f}(\mathbf{x}^*) - \varphi(\mathbf{x}^*)$$

*is a gap function such that, for all $x \in$ dom $\varphi$ and $r(x) \in F(x)$,*

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq l(x).$$

*The function $l$ is convex and weakly lower semi-continuous on dom $\varphi$. Moreover if, in addition, $f$ and $\varphi$ are Lipschitz continuous on bounded subsets of $E$, then $l$ is also Lipschitz continuous on bounded subsets of $E$.*

**Proof.** For all $x \in$ dom $\varphi, r(x) \in F(x) = \partial f(x)$, we have $f(x^*) \geq f(x) + \langle r(x), x^* - x\rangle$. So, we obtain

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq f(x) - f(x^*) + \varphi(x) - \varphi(x^*) = l(x).$$

The rest of the proof is obvious. $\square$

Note that the Lipschitz continuity condition requested on $\varphi$ (respectively $\varphi$ and $f$) in Proposition 5.2 (respectively Proposition 5.3) is not very restrictive. This fact is discussed in the following lemma and comment.

**Lemma 5.3** *Let $g \in \Gamma_0(H)$ and let $B$ be a bounded subset of int(dom $g$). If $\partial g$ is bounded on $B$, then $g$ is Lipschitz continuous on $B$.*

**Proof.** Let $x, y \in B$. Since $B \subseteq \text{int}(\text{dom } g)$, the subdifferentials $\partial g(x)$ and $\partial g(y)$ are nonempty. Let $s(x) \in \partial g(x)$ and $s(y) \in \partial g(y)$. Then

$$\begin{array}{lll} g(x) - g(y) & \leq \langle s(x), x - y \rangle & \leq \|s(x)\| \, \|x - y\| \\ g(y) - g(x) & \leq \langle s(y), y - x \rangle & \leq \|s(y)\| \, \|y - x\|. \end{array}$$

So $|g(x) - g(y)| \leq L\|x - y\|$ where $L = \sup\{\|s(z)\| \mid z \in B, s(z) \in \partial g(z)\}$. Since $\partial g$ is bounded on $B$, this constant $L$ is finite and thus $g$ is Lipschitz continuous on $B$. $\square$

Moreover, we know that $\partial g$ is always bounded on bounded subsets of int(dom $g$) when $H$ is a finite dimensional space.

**Proposition 5.4** *If $F$ is strongly monotone with modulus $\alpha > 0$ on dom $\varphi$ and $x^*$ denotes the unique solution of problem $(GVIP)$, then*

$$\mathbf{l}(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}^*\|^{\mathbf{2}}$$

*is a gap function such that, for all $x \in$ dom $\varphi$ and $r(x) \in F(x)$,*

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq \alpha \, l(x).$$

*Moreover $l$ is strongly convex, weakly lower semi–continuous on $H$ and Lipschitz continuous on bounded subsets of dom $\varphi$.*

**Proof.** Since $x^*$ is the unique solution of problem $(GVIP)$, it is obvious that $l$ is a gap function. Moreover, $l$ is strongly convex and weakly lower semi–continuous on $H$ and for all $x \in$ dom $\varphi$, we have

$$\begin{aligned} &\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \\ &= \langle r(x) - r(x^*), x - x^* \rangle + \langle r(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*). \end{aligned}$$

Since $F$ is strongly monotone with modulus $\alpha$ and $x^*$ is the solution of problem $(GVIP)$, we obtain immediately that the right–hand side of the previous equality is greater than $\alpha\, l(x)$. Finally, let $B$ be a bounded subset of dom $\varphi$. Then there exists $c_1 > 0$ such that $\|z\| \leq c_1$ for all $z \in B$. So, for all $x, y \in B$, we have successively

$$\begin{aligned} \|x - x^*\|^2 - \|y - x^*\|^2 &= \|x - y\|^2 + 2\langle x - y, y - x^*\rangle \\ &\leq \|x - y\| \,[\; \|x - y\| + 2\|y - x^*\| \;] \\ &\leq \|x - y\| \,[\; 4c_1 + 2\|x^*\| \;] \end{aligned}$$

i.e., $l$ is Lipschitz continuous on $B$. □

The condition here below gathers the properties requested on the gap function. These properties are satisfied in the three situations described in Propositions 5.2, 5.3 and 5.4.

**Condition $(GZH_{x^*})$:**

**(i)** *there exist a constant $\alpha > 0$ and a function $l$ defined on dom $\varphi$ such that,*

$$\langle r(y), y - x^*\rangle + \varphi(y) - \varphi(x^*) \geq \alpha\, l(y), \;\; \forall y \in dom\ \varphi, \;\; \forall r(y) \in F(y);$$

**(ii)** *$l(x) \geq 0$ for all $x \in$ dom $\varphi$, and $l(\bar{x}) = 0 \Leftrightarrow \bar{x}$ is a solution of $(GVIP)$;*

**(iii)** *$l$ is weakly lower semi–continuous on $E$ and Lipschitz continuous on bounded subsets of $E$.*

If $\varphi$ is the indicator function of a closed convex set, then Condition $(GZH_{x^*})$ is similar to Condition $(ZH_{x^*})$ imposed by Zhu (see Chapter 2, Section 2.2.2).

The purpose of the next proposition is to show in which cases conditions $(b)$ and $(c)$ of Lemma 5.2 are satisfied.

**Proposition 5.5** *(a) Assume that assumptions of Theorem 5.1 are satisfied as also Condition $(GZH_{x^*})$ (i) and (ii). If $F$ is bounded on bounded subsets of $E$, then $\sum_{k=0}^{+\infty} \mu_k\, l(x^k) < +\infty$.*
*(b) If there exists $c > 0$ such that*

$$\varphi^k(x^k) - \varphi^k(x^{k+1}) \leq c\|x^{k+1} - x^k\| \text{ for all } k, \tag{5.21}$$

*then there exists $\delta > 0$ such that, for all $k$, $\|x^{k+1} - x^k\| \leq \delta\mu_k$.*

**Proof.** (a) Since the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded and $F$ is bounded on bounded subsets of $E$, the sequence $\{r(x^k)\}_{k\in\mathbb{N}}$ is bounded and also the sequence $\{\eta_k\}_{k\in\mathbb{N}}$. Then, using successively Theorem 5.1 and Condition $(GZH_{x^*})$ (i) and (ii), we have

$$\sum_{k=1}^{+\infty}\mu_k[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] < +\infty \text{ and } \sum_{k=1}^{+\infty}\mu_k\, l(x^k) < +\infty.$$

(b) From the optimality conditions (5.5) applied to $x = x^k$, we obtain

$$\begin{aligned}&\langle \nabla K(x^{k+1}) - \nabla K(x^k), x^{k+1} - x^k\rangle \\ &\le (\mu_k/\eta_k)[\langle r(x^k), x^k - x^{k+1}\rangle + \varphi^k(x^k) - \varphi^k(x^{k+1})].\end{aligned}$$

Since $K$ is strongly convex and $\|r(x^k)\| \le \eta_k$, we derive that

$$\beta\|x^{k+1} - x^k\|^2 \le \mu_k\|x^{k+1} - x^k\| + (\mu_k/\eta_k)[\varphi^k(x^k) - \varphi^k(x^{k+1})].$$

Since $\eta_k \ge 1$, we deduce from (5.21) that $\|x^{k+1} - x^k\| \le \delta\mu_k$ for all $k$, with $\delta = (1/\beta)[\,1 + c\,]$. □

**Remark 5.1** If $\varphi^k$ is the indicator function of a closed convex subset $C^k$ of $C$ such that $C^k \subset C^{k+1} \subset C$ for all $k \in \mathbb{N}$, then inequality (5.21) is satisfied.
When $\varphi$ is the indicator function of the set $C = \{x \mid g_i(x) \le 0, i = 1, \ldots, m\}$ with $g_1, \ldots, g_m$ convex functions from $\mathbb{R}^n$ to $\mathbb{R}$, inequality (5.21) is satisfied by the logarithmic barrier functions and by the inverse barrier functions provided that the barrier parameters be large enough.
Indeed, if $\varphi^k(x) = \nu_k^{-1}b(x)$ with

$$b(x) = -\sum_{i=1}^{m}\ln\left(\min\left(\frac{1}{2}, -g_i(x)\right)\right) \quad \text{or} \quad b(x) = -\sum_{i=1}^{m}\frac{1}{g_i(x)}.$$

Let $e^k \in \partial b(x^k)$ (it exists because $b$ is convex and $x^k \in \operatorname{int} C$). Then

$$\begin{aligned}\varphi^k(x^k) - \varphi^k(x^{k+1}) &\le \nu_k^{-1}\langle e^k, x^k - x^{k+1}\rangle \\ &\le \nu_k^{-1}\|e^k\|\,\|x^k - x^{k+1}\|.\end{aligned}$$

So, if $\nu_k \ge \|e^k\|$, then $\varphi^k(x^k) - \varphi^k(x^{k+1}) \le \|x^{k+1} - x^k\|$.

Notice that the choice $\nu_k \ge \|e^k\|$ is possible because $\varphi^k$ is built once $x^k$ is known.

We are now ready to state the main convergence result.

**Theorem 5.3** *Suppose that the following conditions are satisfied:*
*(a) Assumptions of Lemma 5.1 hold;*
*(b) Condition* $(GZH_{x^*})$ *holds;*
*(c)* $F$ *is bounded on bounded subsets of* $E$*;*
*(d) Inequality (5.21) holds;*
*(e)* $\sum_{k=0}^{+\infty} \mu_k = +\infty,\ \sum_{k=0}^{+\infty} \mu_k^2 < +\infty.$
*Then the sequence* $\{x_k\}_{k\in \mathbb{N}}$ *is bounded,* $l(x^k) \to 0$ *and any weak limit point of* $\{x_k\}_{k\in \mathbb{N}}$ *is a solution of problem* $(GVIP)$*.*
*If, in addition, condition (iv) of Lemma 5.1 is satisfied for each solution of problem* $(GVIP)$ *and* $\nabla K$ *is weakly continuous on dom* $\varphi$*, then* $x^k \rightharpoonup \bar{x}$ *and* $\bar{x}$ *is a solution of* $(GVIP)$*. If, in addition, the gap function* $l$ *is strongly convex on an open set containing dom* $\varphi$*, then* $x^k \to x^*$*, the unique solution of* $(GVIP)$*.*

**Proof.** The first part of the theorem follows immediately from Theorem 5.1, Proposition 5.5, Lemma 5.2 and Proposition 5.1.

Suppose now that the sequence $\{x^k\}_{k\in \mathbb{N}}$ has two different weak limit points $x^1$ and $x^2$. Let $\{x^{m(k)}\}_{k\in \mathbb{N}}$ be the subsequence of $\{x^k\}_{k\in \mathbb{N}}$ weakly converging to $x^1$ and $\{x^{n(k)}\}_{k\in \mathbb{N}}$ be the subsequence weakly converging to $x^2$. By the first part of the theorem, $x^1$ and $x^2$ are solutions of problem $(GVIP)$. Then, by Theorem 5.1, the sequences of Lyapunov functions $\{\Gamma^k(x^1, x^k)\}_{k\in \mathbb{N}}$ and $\{\Gamma^k(x^2, x^k)\}_{k\in \mathbb{N}}$ are convergent in $\mathbb{R}$. We denote respectively by $\Gamma_1$ and $\Gamma_2$ their limits. By definition of the Lyapunov function, we have

$$\Gamma^{n(k)}(x^1, x^{n(k)}) - \Gamma^{n(k)}(x^2, x^{n(k)})$$

$$= K(x^1) - K(x^2) - \langle \nabla K(x^{n(k)}), x^1 - x^2\rangle$$

$$+(\mu_{n(k)}/\eta_{n(k)})[\langle r(x^1), x^{n(k)} - x^1\rangle - \langle r(x^2), x^{n(k)} - x^2\rangle + \varphi(x^2) - \varphi(x^1)].$$

If $\nabla K$ is weakly continuous on dom $\varphi$, since $\eta_k \geq 1$ for all $k$ and $\mu_k \to 0$, we obtain, taking the limit on $k$ in the last equality, that

$$\Gamma_1 - \Gamma_2 = K(x^1) - K(x^2) - \langle \nabla K(x^2), x^1 - x^2\rangle. \tag{5.22}$$

Since the role of $x^1$ and $x^2$ is symmetric, we also have that

$$\Gamma_1 - \Gamma_2 = K(x^1) - K(x^2) - \langle \nabla K(x^1), x^1 - x^2\rangle. \tag{5.23}$$

Comparing (5.22) and (5.23), we obtain $\langle \nabla K(x^1) - \nabla K(x^2), x^1 - x^2 \rangle = 0$. Since $\nabla K$ is strongly monotone, this implies that $x^1 = x^2$. So the sequence $\{x^k\}_{k \in I\!N}$ weakly converges to a solution of $(GVIP)$.

If the gap function $l$ is strongly convex with constant $s > 0$ on an open convex set containing dom $\varphi$, then $x^*$ is the unique solution of problem $(GVIP)$, $\partial l(x^*)$ is nonempty, and for any $e^* \in \partial l(x^*)$,

$$l(x^k) - l(x^*) - \langle e^*, x^k - x^* \rangle \geq (s/2)\|x^k - x^*\|^2. \tag{5.24}$$

Since $l(x^k) \to 0$, $l(x^*) = 0$ and $x^k \rightharpoonup x^*$, we obtain, passing to the limit in (5.24) that $\|x^k - x^*\| \to 0$ i.e., $x^k \to x^*$ strongly. This completes the proof. □

For example, if $K(x) = (1/2)x^T x$, for all $x \in H$, then $\nabla K$ is weakly continuous. Note that assumption $(ii)$ of Lemma 5.1 ensures the continuity of $\nabla K$ in the strong topology. So, in a finite dimensional space, the convergence of the whole sequence toward a solution is established without any further assumptions.

**Remark 5.2** Our analysis can be applied directly to the optimization problem $(OP)$ by taking for $F$ the subdifferential of any finite-valued convex and continuous function $f$. In that case, $F$ is paramonotone and Condition $(GZH_{x^*})$ $(i)$ and $(ii)$ is satisfied. If, for all $x \in H$, $K(x) = (1/2)x^T x$, and for all $k \in I\!N$, $\varphi^k = \varphi = \Psi_C$, where $\Psi_C$ denotes the indicator function of a closed convex subset $C$ of $H$, then the method defined by subproblems $(PSAP^k)$, $k \in I\!N$, reduces to the projected subgradient process and our results generalize well-known ones (see, for example, [4]).

**Remark 5.3** If we take $\eta_k = 1$ for all $k$ in subproblems $(PSAP^k)$, the same conclusions as in the preceding theorems can be obtained provided that the following additional condition holds:

**Condition $(C)$:**

$$\exists\, a, b > 0 \ : \ \|r(x)\| \leq a\|x\| + b, \ \forall x \in H, \ \forall r(x) \in F(x).$$

This essentially means that the norm of $F$ does not increase faster than linearly with the norm of $x$. Condition $(C)$ was imposed by Cohen in the nonperturbed setting (see [33]). Observe that this condition implies boundedness of $F$ on bounded subsets and is obviously satisfied when $F$ is Lipschitz continuous and there exists $\bar{y} \in \text{dom } F$ such that $\|F(\bar{y})\|$ is bounded.

To prove this fact, we need the following lemma:

**Lemma 5.4** (See [34], Lemma 5) *Let $\{u^k\}_{k\in\mathbb{N}}$, $\{t^k\}_{k\in\mathbb{N}}$, $\{\delta^k\}_{k\in\mathbb{N}}$ be sequences in $\mathbb{R}^+$ such that*

$$\textstyle\sum_{k=0}^{+\infty} t^k < +\infty,\ \exists\, \delta > 0 : \delta^k \leq \delta\ \forall k \in \mathbb{N},\ \text{ and } u^k \leq \sum_{i=0}^{k-1} t^i\, U^{i+1} + \delta^k,$$

*where $U^k$ denotes $\sup_{i\leq k} u^i$, then the sequence $\{u^k\}_{k\in\mathbb{N}}$ is bounded.*

Let us examine the proof of Lemma 5.1 in the case where $\eta_k = 1$ for all $k$ (so that $\mu_k = \lambda_k$ for all $k$). The terms $s_1, s_{21}$ and $s_3$ of inequality (5.7) can be treated similarly. For the term $s_{22}$, inequalities (5.11) and (5.12) are no more valid. They can be replaced in the following way. From Condition $(C)$ and the fact that $\lambda_k \leq \lambda_0$ for all $k$, we have

$$\begin{aligned}
\lambda_k\langle r(x^k), w^k - x^*\rangle &\leq \lambda_k(a\|x^k\| + b)\,\|w^k - x^*\| \\
&\leq \lambda_k\, a\,\|x^k - x^*\|\,\|w^k - x^*\| \\
&\quad +\lambda_k(a\|x^*\| + b)\,\|w^k - x^*\| \\
&\leq (\lambda_k^2 a^2/(2\theta))\,\|x^k - x^*\|^2 + (\theta/2)\,\|w^k - x^*\|^2 \\
&\quad +\lambda_0(a\|x^*\| + b)\,\|w^k - x^*\|,
\end{aligned} \tag{5.25}$$

$$\begin{aligned}
\lambda_k\langle r(x^k), x^k - x^{k+1}\rangle &\leq (\lambda_k^2/(2\gamma))\,\|r(x^k)\|^2 + (\gamma/2)\,\|x^{k+1} - x^k\|^2 \\
&\leq (\lambda_k^2 a^2/\gamma)\,\|x^k - x^*\|^2 + (\lambda_k^2/\gamma)\,(a\|x^*\| + b)^2 \\
&\quad +(\gamma/2)\,\|x^{k+1} - x^k\|^2,
\end{aligned} \tag{5.26}$$

where the last inequalities hold for any $\theta$, $\gamma > 0$.
So, with (5.11) and (5.12) replaced by (5.25) and (5.26) respectively, inequality (5.3) becomes

$$\begin{aligned}
&\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq -c\|x^{k+1} - x^k\|^2 + \tilde{T}^k + \lambda_k^2\tilde{u} \\
&\quad -\lambda_k[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] + \tilde{c}\lambda_k^2\|x^k - x^*\|^2,
\end{aligned} \tag{5.27}$$

$$\begin{array}{rcl} \text{with} \quad \tilde{c} & = & a^2(1/(2\theta)+1/\gamma), \\ \tilde{T}^k & = & \lambda_0(a\|x^*\|+b)\|w^k-x^*\|+\lambda_0|\varphi^k(w^k)-\varphi(x^*)| \\ & & +(\Lambda^2/(2\tau)+\theta/2)\|w^k-x^*\|^2, \\ \tilde{u} & = & (a\|x^*\|+b)^2/\gamma+(1/(2\mu))\|r(x^*)\|^2, \\ c, & \tau, & \gamma, \text{ and } \mu>0 \text{ as in (5.3).} \end{array}$$

Now, we can show that Theorem 5.1 holds again. Indeed, from inequality (5.14), we derive that

$$\Gamma^{k+1}(x^*,x^{k+1})-\Gamma^k(x^*,x^k)\le \tilde{c}\lambda_k^2\,\|x^k-x^*\|^2+\tilde{T}^k+\tilde{u}\,\lambda_k^2.$$

And if we take the sum in this last inequality, we obtain for all $N\in I\!\!N$ :

$$\Gamma^{N+1}(x^*,x^{N+1})\le\Gamma^0(x^*,x^0)+\tilde{c}\sum_{k=0}^{N}\lambda_k^2\|x^k-x^*\|^2+\sum_{k=0}^{N}(\tilde{T}^k+\tilde{u}\,\lambda_k^2).$$

From relations (5.2) and the convergence of the series $\sum_{k=0}^{+\infty}\lambda_k^2$, it follows that the series $\sum_{k=0}^{+\infty}(\tilde{T}^k+\tilde{u}\,\lambda_k^2)$ is convergent. Then, using inequality (5.6), we can apply Lemma 5.4 and deduce that the sequence $\{x^k\}_{k\in I\!\!N}$ is bounded. The rest of the proof of Theorem 5.1 is obvious.

Now, if the sequence $\{x^k\}_{k\in I\!\!N}$ is known to be bounded, we can say that there exists a constant $e>0$ such that, for all $k\in I\!\!N$, $\|x^k-x^*\|\le e$. Including this in (5.27), we obtain

$$\begin{aligned} \Gamma^{k+1}(x^*,x^{k+1})-\Gamma^k(x^*,x^k)\le{}& -c\|x^{k+1}-x^k\|^2+\tilde{T}^k+\lambda_k^2\tilde{\tilde{u}} \\ & -\lambda_k[\langle r(x^k),x^k-x^*\rangle+\varphi(x^k)-\varphi(x^*)], \end{aligned} \tag{5.28}$$

with $\tilde{\tilde{u}}=\tilde{u}+\tilde{c}e^2$.

Consequently, the proofs of Theorem 5.2, Proposition 5.5 and Theorem 5.3 are obtained simply by using inequality (5.28) instead of (5.3). □

## 5.2 Relaxation of the Scheme and Enlargements of Multivalued Mappings

In this section, we introduce a relaxation in subproblems $(PSAP^k)$ by allowing an inexact computation of an element of $F(x^k)$ in the sense that $r(x^k)$ can be chosen in an enlargement of $F$ at $x^k$. This idea originates from the constrained optimization case:

$$(COP) \quad \min_{x \in C} f(x),$$

where $f$ is a finite–valued convex continuous function defined on $H$ and $C$ is a nonempty closed convex subset of $H$. A well-known method to solve this problem is the classical projected subgradient procedure characterized by the following subproblems:

$$\begin{cases} x^{k+1} = Proj_C\,(x^k - \lambda_k r(x^k)), \\ \text{with } r(x^k) \in \partial f(x^k). \end{cases}$$

It is easy to see that this problem is just subproblem $(SAP^k)$ with the auxiliary function $K$ defined for each $x \in H$ by $K(x) = (1/2)x^T x$. We refer to [4] and the references cited therein for details on convergence of this method. In that paper, the projected subgradient method is relaxed by using elements in the $\epsilon^k$-subdifferential of $f$ at $x^k$. Remind that for $\epsilon \geq 0$, the $\epsilon$-subdifferential of $f$ at $x \in H$ is the set $\partial_\epsilon f(x)$ defined by:

$$\partial_\epsilon f(x) = \{\, u \in H \ : \ f(y) \geq f(x) + < u, y - x > -\epsilon, \ \forall y \in H \,\}.$$

The introduction of the parameter $\epsilon$ produces an enlargement of $\partial f(x)$ with good continuity properties. This generally preserves the convergence properties of the method while giving more latitude and more robustness with respect to numerical errors. This concept is studied, for example, in [16], [20], [59], [60], [95]. It is applied to develop methods of $\epsilon$-descent in [60], bundle methods in [75], [76], [115], [121], [122] or also to devise an inexact proximal point method with generalized Bregman distances in [69].

Following the same idea as in [35], we can see that the projected inexact subgradient algorithm reduces to a projected proximal scheme like that studied in [81] by choosing adequately the sequence $\{\epsilon^k\}_{k \in I\!N}$. Note also that inexact computation of the subgradients in the framework of the projected subgradient method is considered in a different way in [3], [119]. In those papers, the iteration is of the form $x^{k+1} = Proj_C\,(x^k - \lambda_k[r(x^k) + v^k])$ with $r(x^k) \in \partial f(x^k)$ and $\{v^k\}_{k\in I\!N}$ such that $lim_{k\to\infty} v^k = 0$ in [3] or $||v^k|| \leq \tau$ where $\tau$ denotes a maximal error of magnitude in [119].

Similarly, for problem $(GVIP)$, we can choose $r(x^k)$ in an enlargement of $F$ at $x^k$ depending on the iteration $k$. We denote it by $G^k(x^k)$. So, for a given iterate $x^k$, the problem considered at iteration $k$ is now the following:

$$(IPSAP^k)\begin{cases} \text{choose } r^k(x^k) \in G^k(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r^k(x^k) + \lambda_k^{-1}(\nabla K(x^{k+1}) - \nabla K(x^k)), x - x^{k+1}\rangle \\ \qquad\qquad\qquad\qquad +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0. \end{cases}$$

It remains now to describe how to choose these enlargements $G^k$ in such a way that convergence results similar to those of Section 5.1 can be obtained.

When $F$ is monotone, a natural enlargement of $F$ is the $\epsilon$–enlargement introduced in [25]. For $\epsilon \geq 0$, the $\epsilon$-enlargement of the monotone operator $F$ is denoted by $F^\epsilon$ and defined for all $y \in H$ by:

$$F^\epsilon(y) = \{\, u \in H \;:\; \langle r(x) - u, x - y\rangle \geq -\epsilon,\ \forall x \in H,\ \forall r(x) \in F(x)\,\}.$$

It is clear that $F \subset F^0 \subset F^\epsilon$, for all $\epsilon \geq 0$. Moreover, when $F$ is maximal monotone, $F = F^0$. For more details about $F^\epsilon$, we refer the reader to [25], [27]. The $\epsilon$-enlargement is applied in [26] to propose a bundle method to find a zero of a maximal monotone operator. It is also used to devise an inexact proximal point method with Bregman distances for variational inequalities in [25] and to construct a hybrid approximate extragradient-proximal point algorithm in [118]. These works show that when $F$ is maximal monotone, $F^\epsilon$ inherits most properties from the $\epsilon$–subdifferential and plays the role of this one in nonsmooth optimization. One property that $F^\epsilon$ shares with $\partial_\epsilon f$ and that will be used in our analysis is given in the theorem of Brønsted and Rockafellar (see, for example, [20]). For a lower semi–continuous proper convex function $f$, this theorem states that any $\epsilon$–subgradient of $f$ at a point $x^\epsilon$ can be approximated by some exact subgradient computed at some $x$ possibly different from $x^\epsilon$. Here is the generalization of this property for $F^\epsilon$.

**Proposition 5.6** (See [27], Theorem 2.1) *Let $F$ be a maximal monotone operator defined on a Hilbert space $H$, $\epsilon > 0$ and $(x^\epsilon, r^\epsilon) \in$ Graph $F^\epsilon$. Then for all $\eta > 0$, there exists $(x, r) \in$ Graph $F$ such that*

$$\|r - r^\epsilon\| \leq \epsilon/\eta \quad \text{and} \quad \|x - x^\epsilon\| \leq \eta.$$

This formula makes clear that the best compromise will be achieved if we take $\eta = \sqrt{\epsilon}$. However, despite these common behaviors, observe that when $F = \partial f$ for some convex function $f$, we have that $\partial_\epsilon f \subset F^\epsilon$ for all $\epsilon \geq 0$ but we do not get in general that $\partial_\epsilon f = F^\epsilon$. The following examples illustrate this fact and are taken from [25].

**Example 5.1** Let $H = I\!R, f(x) = |x|, F = \partial f$.
In that case, we have

$$\partial_\epsilon f(x) = F^\epsilon(x) = \begin{cases} [\,1 - \epsilon/x, 1\,] & \text{if } x > \epsilon/2, \\ [\,-1, 1\,] & \text{if } |x| \leq \epsilon/2, \\ [\,-1, -1 - \epsilon/x\,] & \text{if } x < -\epsilon/2. \end{cases}$$

**Example 5.2** Let $H = I\!R, f(x) = (1/2)x^2, F = \partial f$.
Then we get

$$\partial_\epsilon f(x) = [\,x - \sqrt{2\epsilon}, x + \sqrt{2\epsilon}\,],$$

$$F^\epsilon(x) = [\,x - 2\sqrt{\epsilon}, x + 2\sqrt{\epsilon}\,].$$

So, we have here that $\partial_\epsilon f \subset F^\epsilon = \partial_{2\epsilon} f$.

**Example 5.3** Let $H = I\!R, f(x) = -\ln x, F = \partial f$.
Then for any $x > 0$:

$\partial_\epsilon f(x) = [\,-(1/x)s_1(\epsilon), -(1/x)s_2(\epsilon)\,]$,

where $s_1(\epsilon), s_2(\epsilon)$ are the two roots of the equation $s - 1 - \ln s = \epsilon$,

such that $0 < s_2(\epsilon) \leq 1 \leq s_1(\epsilon)$,

$F^\epsilon(x) = [-(1/x)(1 + \epsilon + 2\sqrt{\epsilon}), (1/x)\min(\epsilon - 1, 0)]$.

We can see that $0 \notin \partial_\epsilon f(x)$ for any $x > 0$ and any $\epsilon \geq 0$ while $0 \in F^\epsilon(x)$ for all $x > 0$ and all $\epsilon \geq 1$. Therefore, if $\epsilon \geq 1$ then $F^\epsilon(x) \not\subset \partial_{\bar{\epsilon}} f(x)$ for any $x > 0$ and any $\bar{\epsilon} \geq 0$.

**Example 5.4** Let $H = I\!R^n, f(x) = x^T Qx + b, Q \in I\!R^{n\times n}$ symmetric and positive definite, $b \in I\!R^n, F = \partial f$.
Then
$$\partial_\epsilon f(x) = \{Qx + b + w : w^T Q^{-1} w \le 2\epsilon\},$$
$$F^\epsilon(x) = \{Qx + b + w : w^T Q^{-1} w \le 4\epsilon\}.$$
Hence, $\partial_\epsilon f \subset F^\epsilon = \partial_{2\epsilon} f$.

Let us come back to our choice of an enlargement of $F$ in our algorithm. If $x^*$ denotes a solution of problem $(GVIP)$, then it is easy to see that, for all $y \in H$ and all $\epsilon \ge 0$:
$$F^\epsilon(y) \subset \{\, u \in H \ : \ \langle u, y - x^*\rangle + \varphi(y) - \varphi(x^*) \ge -\epsilon \,\}.$$
This last set will be denoted by $F^\epsilon_{x^*}(y)$. If we take $G^k \subset F^{\epsilon^k}_{x^*}$ for all $k \in I\!N$, where $\{\epsilon^k\}_{k\in I\!N}$ is a sequence of nonnegative numbers converging to zero, then we can obtain similar results to Lemma 5.1, Theorems 5.1, 5.2. This fact is detailed in next section. Obviously, $F^{\epsilon^k}_{x^*}$ is used only to denote the largest enlargement in which we can take elements. In practice, we will choose a subenlargement $G^k$ independent of a solution of problem $(GVIP)$. For example, we can work with $G^k = F^{\epsilon^k}$ for all $k \in I\!N$. However, we meet some problems to prove the results corresponding to Proposition 5.5 and Theorem 5.3 with $G^k \subset F^{\epsilon^k}_{x^*}$. This would come from the fact that the enlargement $F^{\epsilon^k}_{x^*}$ is too big. To remedy this problem, we will introduce a new enlargement of $F$ contained in $F^{\epsilon^k}_{x^*}$.

Assume that for some solution $x^*$ of problem $(GVIP)$, there exists a constant $\alpha > 0$ and a function $l$ such that the mapping $F$ satisfies Condition $(GZH_{x^*})(i)$ and $(ii)$. The $\alpha$-$\epsilon$-enlargement of $F$ around $x^*$, denoted by $F^{\epsilon,\alpha}_{x^*}$, is defined at $y \in \mathrm{dom}\ \varphi$ by:
$$F^{\epsilon,\alpha}_{x^*}(y) = \{\, u \in H \ : \ \langle u, y - x^*\rangle + \varphi(y) - \varphi(x^*) \ge \alpha\, l(y) \ - \epsilon \,\}.$$
We clearly see that $F \subset F^{\epsilon,\alpha}_{x^*} \subset F^\epsilon_{x^*}$. So to obtain generalizations of Proposition 5.5 and Theorem 5.3, we will choose $G^k \subset F^{\epsilon^k,\alpha}_{x^*}$ for all $k$. Again, in practice, we work with a subenlargement without any reference to a solution $x^*$. In the following examples, we comment somewhat how to choose $G^k$ independently of $x^*$.

**Example 5.5** If $F$ is paramonotone on dom $\varphi$ and $F(x)$ is a bounded and weakly closed subset of $H$ for all $x \in$ dom $\varphi$, then it is shown in Proposition 5.2 that Condition $(GZH_{x^*})$ $(i)$ and $(ii)$ is satisfied for any solution of problem $(GVIP)$ with

$$l(y) = \inf_{r(y)\in F(y)} \langle r(y), y - x^*\rangle + \varphi(y) - \varphi(x^*),\ \forall y \in \mathrm{dom}\varphi,\ \text{and } \alpha = 1.$$

In that case, for all $y \in H$,

$$F^{\epsilon,\alpha}_{x^*}(y) = \{\, u \in H : \inf_{r(y)\in F(y)} \langle r(y) - u, y - x^*\rangle \le \epsilon \,\}.$$

For this example, $G^k = F$, for all $k$, is the only well-known subenlargement of $F^{\epsilon^k,\alpha}_{x^*}$ independent of $x^*$.
Finally, let us illustrate by an example that we can have $F^{\epsilon,\alpha}_{x^*}$ strictly contained in $F^{\epsilon}_{x^*}$. Indeed, easy calculations show that when $F(y) = y$, $\forall y \in I\!R$, and $\varphi = 0$, then $x^* = 0$, and

$$F^{\epsilon,\alpha}_{x^*}(y) = \begin{cases} [y - \epsilon/y, +\infty[ & \text{if } y > 0, \\ ]-\infty, y - \epsilon/y] & \text{if } y < 0, \\ I\!R & \text{if } y = 0; \end{cases}$$
$$F^{\epsilon}_{x^*}(y) = \begin{cases} [-\epsilon/y, +\infty[ & \text{if } y > 0, \\ ]-\infty, -\epsilon/y] & \text{if } y < 0, \\ I\!R & \text{if } y = 0. \end{cases}$$

**Example 5.6** In the important particular case where $F$ is the subdifferential of a finite-valued convex continuous function $f$, then as already said in Proposition 5.3, Condition $(GZH_{x^*})$ $(i)$ and $(ii)$ is verified for any solution $x^*$ of problem $(OP)$ with

$$l(y) = f(y) + \varphi(y) - f(x^*) - \varphi(x^*),\ \forall y \in H,\ \text{and } \alpha = 1.$$

If we write the enlargement $F^{\epsilon,\alpha}_{x^*}$ in this case, we obtain, for all $y \in H$,

$$F^{\epsilon,\alpha}_{x^*}(y) = \{\, u \in H :\ f(x^*) \ge f(y) + \langle u, x^* - y\rangle - \epsilon \,\}.$$

This definition makes us think to the $\epsilon$-subdifferential of $f$ at $y$. Indeed, we see that $\partial_\epsilon f \subset F^{\epsilon,\alpha}_{x^*}$. So, we can take $G^k = \partial_{\epsilon^k} f$, for all $k$, in that case.

**Example 5.7** If $F$ is a strongly monotone operator with modulus $\alpha$, then we know from Proposition 5.4 that Condition $(GZH_{x^*})$ holds for the unique solution $x^*$ of problem $(GVIP)$ with

$$l(y) = \|y - x^*\|^2,\ \forall y \in H.$$

In that case, for all $y \in H$, we have

$$F_{x^*}^{\epsilon,\alpha}(y) = \{\, u \in H : \langle u, y - x^* \rangle + \varphi(y) - \varphi(x^*) \geq \alpha \|y - x^*\|^2 - \epsilon \,\}.$$

If we consider the $\alpha$-$\epsilon$-enlargement of $F$ without any reference to a solution $x^*$, we can define like in [111],

$$F^{\epsilon,\alpha}(y) = \{\, u \in H : \langle u - r(x), y - x \rangle \geq \alpha \|y - x\|^2 - \epsilon,\ \forall x \in H,\ \forall r(x) \in F(x) \,\}.$$

It is easy to see that $F^{\epsilon,\alpha} \subset F_{x^*}^{\epsilon,\alpha}$. So, we could choose $G^k = F^{\epsilon^k,\alpha}$ for all $k$. As an illustration, let us consider the linear mapping defined by $F(y) = Qy + b,\ \forall y \in I\!R^n$, with $Q \in I\!R^{n\times n}$ a symmetric positive definite matrix, and $b \in I\!R^n$. The mapping $F$ is strongly monotone with modulus $\alpha = \lambda_{min}(Q) > 0$ where $\lambda_{min}(Q)$ denotes the minimum eigenvalue of $Q$. Explicit computation of $F^\epsilon(y)$ and $F^{\epsilon,\alpha}(y)$ for any $y \in I\!R^n$, gives:

$$\begin{aligned} F^\epsilon(y) &= \{\, Qy + b + w : \ w^T Q^{-1} w \leq 4\epsilon \,\}; \\ F^{\epsilon,\alpha}(y) &= \{\, Qy + b + w : \ w^T [Q - \alpha I]^{-1} w \leq 4\epsilon \,\}. \end{aligned} \tag{5.29}$$

Let us justify the expression of $F^{\epsilon,\alpha}$ for this example. Let $y \in I\!R^n$. By definition, $u \in F^{\epsilon,\alpha}(y)$ if and only if for all $x \in I\!R^n$,

$$\langle u, x - y \rangle \leq \langle Qx + b, x - y \rangle - \alpha \|y - x\|^2 + \epsilon.$$

Denoting $x - y = \lambda d$, with $\lambda > 0$ and $d \in I\!R^n$, we obtain that $u \in F^{\epsilon,\alpha}(y)$ if and only if for all $d \in I\!R^n$,

$$\begin{aligned} \langle u, d \rangle &\leq \inf_{\lambda > 0} \{ \langle Qy + b, d \rangle + \lambda [\langle Qd, d \rangle - \alpha \|d\|^2] + \epsilon/\lambda \} \\ &= \langle Qy + b, d \rangle + \sqrt{4\epsilon\, d^T (Q - \alpha I) d}. \end{aligned}$$

Moreover, we can verify that

$$\sqrt{4\epsilon\, d^T (Q - \alpha I) d} = \sup_{w \in E} \langle w, d \rangle,$$

with $E = \{\, w \in I\!R^n : \ w^T (Q - \alpha I)^{-1} w \leq 4\epsilon \,\}$.
Then we deduce (5.29) from [60] (Chapter 3, Theorem 4.1.1). □

## 5.3 Convergence Results for the Inexact Perturbed Auxiliary Problem Method

In this section, we generalize the results presented in Section 5.1 to the relaxed scheme described by the following subproblems:

$$(IPSAP^k)\begin{cases} \text{choose } r^k(x^k) \in G^k(x^k) \text{ and} \\ \text{find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r^k(x^k) + \lambda_k^{-1}(\nabla K(x^{k+1}) - \nabla K(x^k)), x - x^{k+1}\rangle \\ \qquad\qquad\qquad\qquad +\varphi^k(x) - \varphi^k(x^{k+1}) \geq 0, \end{cases}$$

where $\{G^k\}_{k\in I\!N}$ is a sequence of multivalued mappings defined onto $H$ such that $F \subset G^k$ for all $k \in I\!N$ and the sequence $\{\lambda_k\}_{k\in I\!N}$ takes the form:

$$\begin{cases} \lambda_k = \mu_k/\eta_k, \forall k \in I\!N, \text{ with } \{\mu_k\}_{k\in I\!N} \text{ a sequence of positive numbers,} \\ \text{and } \eta_k = \begin{cases} \max\{1, \|r^0(x^0)\|\}, \text{ if } k = 0, \\ \max\{\eta_{k-1}, \|r^k(x^k)\|\}, \text{ if } k \geq 1. \end{cases} \end{cases}$$

The generalization of Lemma 5.1 holds for any enlargement $G^k$ of $F$. Generalizations of Theorems 5.1 and 5.2 deal with $G^k \subset F^{\epsilon^k}_{x^*}$ while those of Proposition 5.5 and Theorem 5.3 are restricted to $G^k \subset F^{\epsilon^k,\alpha}_{x^*}$. For the sake of completeness, we write entirely the statement of these generalizations.

**Lemma 5.5** *Assume that $F$ is a monotone multivalued mapping defined on $H$, that problem $(GVIP)$ admits at least one solution denoted by $x^*$, and that the following conditions are satisfied:*

**(i)** $K : H \to I\!R$ *is continuously differentiable and strongly convex with modulus $\beta > 0$ over dom $\varphi$;*

**(ii)** $\nabla K$ *is a Lipschitz continuous mapping with Lipschitz constant $\Lambda$ over dom $\varphi$;*

**(iii)** $\{\mu_k\}_{k\in I\!N}$ *is a nonincreasing sequence of positive numbers;*

**(iv)** $\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ *are such that $\varphi \leq \varphi^k$ for all $k$, and there exists a sequence $\{w^k\}_{k\in I\!N}$ such that*

$$\sum_{k=0}^{+\infty} \|w^k - x^*\| < +\infty \quad \textit{and} \quad \sum_{k=0}^{+\infty} |\varphi^k(w^k) - \varphi(x^*)| < +\infty. \tag{5.30}$$

*Then, if $\{x^k\}_{k\in\mathbb{N}}$ denotes the sequence generated by solving subproblems $(IPSAP^k)$, we have for all $k \in \mathbb{N}$,*

$$\begin{aligned}\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \le -c\|x^{k+1} - x^k\|^2 + T^k + \mu_k^2\, u \\ -(\mu_k/\eta_k)[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)],\end{aligned} \tag{5.31}$$

*where the Lyapunov function $\Gamma^k$ is defined in (5.1), $c, u > 0$, $T^k \ge 0$, and $\sum_{k=0}^{+\infty} T^k < +\infty$.*

**Proof.** Same proof as for Lemma 5.1. □

The following theorem analyzes boundedness of the sequence when we take $G^k \subset F_{x^*}^{\epsilon^k}$ for all $k$.

**Theorem 5.4** *Assume that all assumptions of Lemma 5.5 hold.*
*If $\sum_{k=0}^{+\infty} \mu_k^2 < +\infty$ and there exists $\epsilon > 0$ such that $0 \le \epsilon^k \le \mu_k \epsilon$ for all $k \in \mathbb{N}$, then provided that $x^0 \in dom\ \varphi$, the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by solving subproblems $(IPSAP^k)$ with $G^k \subset F_{x^*}^{\epsilon^k}$ is bounded. Moreover,*

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty, \ \textit{and}$$

$$\sum_{k=0}^{+\infty} (\mu_k/\eta_k)[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] < +\infty.$$

**Proof.** Since $r^k(x^k) \in F_{x^*}^{\epsilon^k}(x^k)$, $\epsilon^k \le \epsilon\mu_k$, and $\eta_k \ge 1$ for all $k$, we have that

$$(\mu_k/\eta_k)[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] \ge -\epsilon^k \mu_k/\eta_k \ge -\epsilon\, \mu_k^2. \tag{5.32}$$

So, we derive from (5.31) that

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \le T^k + (u + \epsilon)\, \mu_k^2. \tag{5.33}$$

Since the series $\sum_{k=0}^{+\infty} T^k$, $\sum_{k=0}^{+\infty} \mu_k^2$ are convergent, inequality (5.33) ensures that $\{\Gamma^k(x^*, x^k)\}_{k\in\mathbb{N}}$ is a Cauchy sequence and thus converges in $H$. Using inequality (5.6), we can conclude that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded. The rest of the proof is the same as for Theorem 5.1 □

To prove that at least one weak limit point of the sequence $\{x^k\}_{k\in\mathbb{N}}$

is a solution of problem $(GVIP)$, we have to impose that $F$ is paramonotone and that $F$ and the operators $\{G^k\}_{k\in\mathbb{N}}$ satisfy the following condition:

**Condition (I):**
*For any bounded sequence $\{z^k\}_{k\in\mathbb{N}}$ of $E$, and $\{r^k\}_{k\in\mathbb{N}}$ with $r^k \in G^k(z^k)$,*

**(I.1)** *the sequence $\{r^k\}_{k\in\mathbb{N}}$ is bounded, and*
**(I.2)** *there exist subsequences $\{z^{k'}\}_{k'\in K\subset\mathbb{N}}$ and $\{r^{k'}\}_{k'\in K\subset\mathbb{N}}$ such that $z^{k'} \rightharpoonup \bar{z}$, $r^{k'} \rightharpoonup \bar{r}, \bar{r} \in F(\bar{z})$, and $\underline{\lim}_{k'\to\infty}\langle r^{k'}, z^{k'} - \bar{z}\rangle + \varphi(z^{k'}) - \varphi(\bar{z}) \geq 0$.*

Obviously, if $G^k = F$ for all $k$, $F$ is monotone, bounded on bounded subsets of dom $\varphi$ and weakly closed on dom $\varphi$, then Condition $(I)$ is satisfied. Observe that if $G^k = F^{\epsilon^k}$ and $\epsilon^k \leq \bar{\epsilon}$ for all $k$, then $G^k \subset F^{\bar{\epsilon}}$ for all $k$. When $F$ is maximal monotone and dom $F$ is closed, $F^{\bar{\epsilon}}$ is locally bounded in the interior of its domain (see [27]). In finite dimension, if $F$ is maximal monotone and dom $F$ is closed, Condition $(I.1)$ is satisfied and Condition $(I.2)$ holds when $\epsilon^k \to 0$ (see [25]). A typical example is that of the subdifferential of a lower semi–continuous proper convex function $f$ in $\mathbb{R}^n$ (see [107]).

**Theorem 5.5** *Suppose that all assumptions of Lemma 5.5 are satisfied and that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded. If $\lim_{k\to\infty}\mu_k = 0$, $\sum_{k=0}^{+\infty}\mu_k = +\infty$, $\lim_{k\to\infty}\epsilon^k = 0$, $F$ is paramonotone over dom $\varphi$, $F$ and $\{G^k\}_{k\in\mathbb{N}}$ satisfy Condition $(I)$, then the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by solving subproblems $(IPSAP^k)$ with $G^k \subset F^{\epsilon^k}_{x^*}$, has at least one weak limit point which is a solution of problem $(GVIP)$.*

**Proof.** The proof is similar to that of Theorem 5.2 with "$r(x^k) \in F(x^k)$" replaced by "$r^k(x^k) \in F^{\epsilon^k}_{x^*}(x^k)$". It suffices to rewrite that proof by using Condition $(I.1)$ instead of "$F$ bounded on bounded subsets of dom $\varphi$" and Condition $(I.2)$ instead of "$F$ weakly closed on dom $\varphi$". □

Now, to ensure that each weak limit point of $\{x^k\}_{k\in\mathbb{N}}$ is a solution of problem $(GVIP)$, we will use a gap function for problem $(GVIP)$ (and thus Condition $(GZH_{x^*})$) as in Section 5.1. In view of Proposition 5.1 and Lemma 5.2, we have to show that Proposition 5.5 can be generalized to the inexact case. For that purpose, we use $F^{\epsilon^k,\alpha}_{x^*}$, the $\alpha$-$\epsilon^k$-enlargement of $F$ around $x^*$, instead of $F^{\epsilon^k}_{x^*}$ in subproblems $(IPSAP^k)$. More precisely, we will choose $r^k(x^k)$ in $G^k(x^k)$ with the operator $G^k$ taken such that $G^k \subset F^{\epsilon^k,\alpha}_{x^*}$.

On the other hand, when $F$ is Lipschitz continuous, we can use the enlargement $F^{\epsilon^k}$.

**Proposition 5.7** *(a) Assume that assumptions of Theorem 5.4 are satisfied as also Condition* $(GZH_{x^*})$ $(i)$ *and* $(ii)$. *If* $G^k$ *satisfies Condition* $(I.1)$ *and either* **(case 1)** $G^k \subset F^{\epsilon^k,\alpha}_{x^*}$, $\forall k$, *and* $\exists\, \epsilon > 0 \,:\, 0 \le \epsilon^k \le \mu_k\, \epsilon$, $\forall k$,

*or* **(case 2)** $F$ *is maximal monotone, Lipschitz continuous on* $H$,

$$G^k \subset F^{\epsilon^k},\ \forall k,\ \textit{and}\ \exists\, \epsilon > 0 \,:\, 0 \le \epsilon^k \le \mu_k^2\, \epsilon^2,\ \forall k,$$

*then* $\sum_{k=0}^{+\infty} \mu_k\, l(x^k) < +\infty$.

*(b) If there exists* $c > 0$ *such that (5.21) holds, then there exists* $\delta > 0$ *such that, for all* $k$, $\|x^{k+1} - x^k\| \le \delta\mu_k$.

**Proof.** (a) First, we prove that, for each $k$, there exists $v^k > 0$ such that

$$\langle r^k(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*) \ge \alpha\, l(x^k) - v^k. \tag{5.34}$$

This is true in **case 1** when $G^k \subset F^{\epsilon^k,\alpha}_{x^*}$ for all $k$. Indeed, $r^k(x^k) \in F^{\epsilon^k,\alpha}_{x^*}(x^k)$ means that (5.34) is satisfied with $v^k = \epsilon^k$. Now, consider **case 2** where $F$ is maximal monotone, Lipschitz continuous on $H$ and $G^k \subset F^{\epsilon^k}$ for all $k$. Since $F$ satisfies Condition $(GZH_{x^*})$ $(i)$ and $(ii)$, we have that for any $r(x^k) \in F(x^k)$,

$$\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*) \ge \alpha\, l(x^k).$$

Therefore, we can write for any $r(x^k) \in F(x^k)$,

$$\begin{aligned} &\langle r^k(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*) \\ &\ge \alpha\, l(x^k) + \langle r^k(x^k) - r(x^k), x^k - x^* \rangle. \end{aligned} \tag{5.35}$$

Moreover, since $r^k(x^k) \in F^{\epsilon^k}(x^k)$ and $F$ is maximal monotone, an extension of the Brønsted and Rockafellar's Theorem (see Proposition 5.6) ensures that, for any $k$:

$$\begin{aligned} &\exists\, (y^k, r(y^k)) \in \text{ Graph } F \text{ such that} \\ &\|x^k - y^k\| \le \sqrt{\epsilon^k} \text{ and } \|r^k(x^k) - r(y^k)\| \le \sqrt{\epsilon^k}. \end{aligned} \tag{5.36}$$

Moreover, since $F$ is Lipschitz continuous on $H$, there exists $L > 0$ such that $e(F(y^k), F(x^k)) \leq L\|y^k - x^k\|$ so that we have that there exists $r^k \in F(x^k)$ such that

$$\|r(y^k) - r^k\| \leq L\|y^k - x^k\| + \sqrt{\epsilon^k}. \tag{5.37}$$

Hence,

$$\begin{aligned}
&\langle r^k(x^k) - r^k, x^* - x^k\rangle \\
&\qquad = \langle r^k(x^k) - r(y^k), x^* - x^k\rangle + \langle r(y^k) - r^k, x^* - x^k\rangle \\
&\qquad \leq \sqrt{\epsilon^k}\,(2 + L)\,\|x^k - x^*\|,
\end{aligned} \tag{5.38}$$

where the inequality comes from (5.36) and (5.37).
The sequence $\{x^k\}_{k\in I\!N}$ being bounded, there exists a constant $e > 0$ such that $||x^k - x^*|| \leq e$ for all $k$. Hence, combining (5.35) with $r(x^k) = r^k$ and (5.38), we deduce that

$$\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*) \geq \alpha\, l(x^k) - \sqrt{\epsilon^k}\,(2 + L)\,e,$$

i.e. (5.34) is satisfied with $v^k = \sqrt{\epsilon^k}(2 + L)e$.
So, if we couple the two situations, we obtain (5.34) with

$$v^k = \begin{cases} \epsilon^k, & \text{in case 1;} \\ \\ \sqrt{\epsilon^k}\,(1 + L)\,e, & \text{in case 2.} \end{cases}$$

We can then deduce that

$$\begin{aligned}
&(\alpha/\bar{\eta})\textstyle\sum_{k=0}^{+\infty} \mu_k\, l(x^k) \\
&\leq \alpha \textstyle\sum_{k=0}^{+\infty}(\mu_k/\eta^k)\, l(x^k) \\
&\leq \textstyle\sum_{k=0}^{+\infty}(\mu_k/\eta^k)\,[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] + \sum_{k=0}^{+\infty}(\mu_k/\eta^k)\, v^k \\
&\leq \textstyle\sum_{k=0}^{+\infty}(\mu_k/\eta^k)\,[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] + \sum_{k=0}^{+\infty} \mu_k\, v^k,
\end{aligned}$$

where the first inequality follows from the fact that $\{x^k\}_{k\in I\!N}$ is bounded and consequently the sequences $\{r^k(x^k)\}_{k\in I\!N}$ and $\{\eta_k\}_{k\in I\!N}$ are bounded, the second inequality comes from (5.34) and the third one follows from the fact that $\eta_k \geq 1$ for all $k$.
From Theorem 5.4, we have that

$$\sum_{k=0}^{+\infty}(\mu_k/\eta^k)\,[\langle r^k(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] < +\infty.$$

Moreover, the bounds imposed on $\epsilon^k$ and the convergence of $\sum_{k=0}^{+\infty} \mu_k^2$ ensure that $\sum_{k=0}^{+\infty} \mu_k \, v^k < +\infty$ and thus that $\sum_{k=0}^{+\infty} \mu_k \, l(x^k) < +\infty$.
(b) Same proof as for Proposition 5.5(b). $\square$

The generalization of Theorem 5.3 can then be deduced.

**Theorem 5.6** *Suppose that the following conditions are satisfied:*
*(a) Assumptions of Lemma 5.5 hold;*
*(d) Condition $(GZH_{x^*})$ holds;*
*(c) $G^k$ satisfies Condition (I.1);*
*(d) Inequality (5.21) holds;*
*(e) $\sum_{k=0}^{+\infty} \mu_k = +\infty, \; \sum_{k=0}^{+\infty} \mu_k^2 < +\infty$;*
*(f) either* **(case 1)** $G^k \subset F_{x^*}^{\epsilon^k,\alpha}$, $\forall k$, *and* $\exists\, \epsilon > 0 \, : \, 0 \le \epsilon^k \le \mu_k \epsilon$, $\forall k$,

*or* **(case 2)** *$F$ is maximal monotone, Lipschitz continuous on $H$, and* $G^k \subset F^{\epsilon^k}$, $\forall k$, *and* $\exists\, \epsilon > 0 \, : \, 0 \le \epsilon^k \le \mu_k^2 \epsilon^2$, $\forall k$.

*Then, the sequence $\{x_k\}_{k \in I\!N}$ is bounded, $l(x^k) \to 0$ and any weak limit point of $\{x_k\}_{k \in I\!N}$ is a solution of problem $(GVIP)$.*
*If, in addition, condition (iv) of Lemma 5.5 is satisfied for each solution of problem $(GVIP)$ and $\nabla K$ is weakly continuous on dom $\varphi$, then $x^k \rightharpoonup \bar{x}$ and $\bar{x}$ is a solution of $(GVIP)$. If, in addition, the gap function $l$ is strongly convex on an open set containing dom $\varphi$, then $x^k \to x^*$, the unique solution of $(GVIP)$.*

**Proof.** The first part of the theorem follows directly from Lemma 5.5, Theorem 5.4, Proposition 5.7, Lemma 5.2 and Proposition 5.1. The rest of the proof is identical to that of Theorem 5.3. $\square$

**Remark 5.4** In the case where $F$ is the subdifferential of a finite–valued convex continuous function $f$ such that problem $(GVIP)$ reduces to the optimization problem $(OP)$, we know that $F$ is paramonotone and satisfies Condition $(GZH_{x^*})$ $(i)$ and $(ii)$. Moreover, as discussed in Section 5.2, we can take for $G^k$ the $\epsilon^k$–subdifferential of $f$. In that case, Condition $(I)$ holds in $I\!R^n$.
Moreover, in the particular case where, for all $x \in H$, $K(x) = (1/2)x^T x$, and for all $k \in I\!N$, $\varphi^k = \varphi = \Psi_C$, where $\Psi_C$ denotes the indicator function of a closed convex subset $C$ of $H$, then the method defined by subproblems

$(IPSAP^k),\ k \in I\!N$, reduces to the projected inexact subgradient process:

$$\begin{cases} x^{k+1} = Proj_C\,[x^k - (\mu_k/\eta_k) r^k(x^k)], \\ \text{with } r^k(x^k) \in \partial_{\epsilon^k} f(x^k), \\ \text{and } \eta_k = \begin{cases} max\{1, \|r^0(x^0)\|\}, \text{ if } k = 0; \\ max\{\eta_{k-1}, \|r^k(x^k)\|\}, \text{ if } k \geq 1. \end{cases} \end{cases}.$$

If problem $(OP)$ is solvable, Theorem 5.6 reduces, for this scheme, to the convergence result of [4].

**Remark 5.5** We can also generalize Remark 5.3 to this inexact case. If we take $\eta_k = 1$ for all $k$ in subproblems $(IPSAP^k)$, the preceding theorems still hold under the following additional condition:

**Condition $(IC)$:**

$$\exists\ a, b > 0\ :\ \forall k \in I\!N\ :\ \|r^k(x)\| \leq a\|x\| + b,\ \forall x \in H,\ \forall r^k(x) \in G^k(x).$$

The proof of this fact is the same as in Remark 5.3.

Observe that Condition $(IC)$ implies Condition $(I.1)$. Note also that Condition $(IC)$ holds when $F$ is maximal monotone, Lipschitz continuous on $H$, there exists $\bar{y} \in \text{dom } F$ such that $F(\bar{y})$ is bounded and $G^k = F^{\epsilon^k}$. To see this, it suffices to use the extension of the Brønsted and Rockafellar's Theorem in the same way as in (5.36).

# Chapter 6

# A Bundle Method to Solve Multivalued Variational Inequalities

In this chapter, we use the perturbations in combination with a bundle strategy to solve problem $(GVIP)$ in the case where

$$\varphi = p + \Psi_C,$$

with $p : H \to I\!R \cup \{+\infty\}$ a lower semi-continuous proper convex function and $\Psi_C$ the indicator function of a nonempty closed convex subset $C$ of $H$ such that $C \subseteq \text{dom } p \subseteq \text{dom } F$ and $C$ is equal to the closure of its interior (so that int $C \neq \emptyset$). The problem considered is thus the following:

$$(GVIP) \begin{cases} \text{find } x^* \in C \text{ and } r(x^*) \in F(x^*) \text{ such that, for all } x \in C, \\ \\ \langle r(x^*), x - x^* \rangle + p(x) - p(x^*) \geq 0. \end{cases}$$

The idea comes from the optimization case where $F = 0$ and $C = H$ such that problem $(GVIP)$ reduces to minimize $p$ on $H$. That problem can be solved by the so–called bundle method, introduced in the eighties by Lemaréchal (see, for example, [35]). In this method, the effective domain of $p$ is supposed to be the whole space $H$ and the strategy is to approximate, at iteration $k$, the function $p$, step by step, by a piecewise linear convex function $p^k$, and to move to the next iterate only when the approximation is suitable. As proven in [35], this method can be seen as a practical implementation of the classical proximal method in convex optimization.

Our purpose in this chapter is to use the bundle strategy to solve the more general problem ($GVIP$). However, there immediately appears a difficulty when we follow this strategy for solving approximately the auxiliary subproblem ($SAP^k$). Indeed, a way to build a piecewise linear convex approximation $p^k \leq p$, is to generate points $y^1, \ldots, y^t$ in $C$, and to consider the function

$$p^k(x) = \max_{1 \leq i \leq t} \{p(y^i) + \langle s(y^i), x - y^i \rangle\}, \quad x \in C, \tag{6.1}$$

where $s(y^i)$ denotes one subgradient of $p$ at $y^i$ for $i = 1, \ldots, t$. Usually the points $y^1, \ldots, y^t$ are the trial points built from $x^k$. Doing that, we suppose that $s(y^i)$ exists for $i = 1, \ldots, t$. But we know that the subdifferential of a convex function is nonempty in the interior of its domain and may be empty on the boundary of its domain (see [106], Theorem 23.4). Here, the latter may occur because $C \subseteq \text{dom}\, p$ and the trial points are in $C$. So, in our method, to prevent the iterates to go to the boundary of $C$, we introduce a barrier function $b(\nu_k, \cdot)$ in the objective function of subproblem ($SAP^k$). Then this problem becomes an unconstrained problem.

In the sequel, we first set the conditions to be satisfied by the approximations $p^k$ of $p$ and we show how to build, step by step, a suitable piecewise linear approximation $p^k$ by means of a bundle strategy. Then, in a second part, we prove the convergence of the general algorithm. The results of this chapter are presented in [114].

## 6.1 Approximate Auxiliary Subproblem and Bundle Strategy

In this section, we first set the approximate auxiliary subproblem that will be considered at each iteration. Then, we propose a bundle scheme to build an approximation of $p$. We show that under some conditions, this scheme produces a suitable approximation after finitely many iterations and we give some classical examples where the conditions are satisfied.

Consider at iteration $k$ the symmetric auxiliary subproblem with a fixed auxiliary function $K$ recalled here below:

$$(SAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and find } x^{k+1} \text{ the solution of} \\ \min_{x\in C}\{p(x) + \lambda_k^{-1}[\ K(x) - K(x^k) - \langle z^k, x - x^k\rangle\ ]\ \}, \\ \text{where } z^k = \nabla K(x^k) - \lambda_k r(x^k). \end{cases}$$

We approximate this problem by replacing the functions $p$ and $\Psi_C$ by functions $p^k$ and $b(\nu_k, \cdot)$ in $\Gamma_0(H)$ respectively, so that we obtain the following subproblem:

$$(BSAP^k)\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and find } x^{k+1} \text{ the solution of} \\ \min_{x\in H}\{p^k(x) + b(\nu_k, x) + \lambda_k^{-1}[K(x) - K(x^k) - \langle z^k, x - x^k\rangle]\}, \\ \text{where } z^k = \nabla K(x^k) - \lambda_k r(x^k). \end{cases}$$

Recall that since $K$ is strongly convex on $C$, there exists one and only one solution for problem $(BSAP^k)$.

The function $\Psi_C$ is approximated by a sequence of barrier functions $\{b(\nu_k, \cdot)\}_k$ associated with $C$ where $\{\nu_k\}_{k\in I\!N}$ is the sequence of positive barrier parameters strictly increasing to $+\infty$. Recall that for all $k$, $b(\nu_k, \cdot)$ is continuous and positive on the interior of $C$ and takes the value $+\infty$ elsewhere. Moreover, for each $x$ in the interior of $C$, the sequence $\{b(\nu_k, x)\}_k$ is strictly decreasing to zero. We refer to Examples 3.3 and 3.7 for some instances and convergence properties. The use of a barrier function in subproblem $(BSAP^k)$ ensures that the iterate $x^{k+1}$ belongs to int $C$. In a similar way as in the preceding chapter, we impose the following condition on the sequence $\{b(\nu_k, \cdot)\}_k$: for each solution $x^*$ of problem $(GVIP)$, there exists a sequence $\{w^k\}_{k\in I\!N}$ in int $C$ such that

$$\sum_{k=1}^{+\infty} \|w^k - x^*\| < +\infty \quad \text{and} \quad \sum_{k=0}^{+\infty} b(\nu_k, w^k) < +\infty. \tag{6.2}$$

Now, for the function $p$, we consider a sequence $\{p^k\}_{k\in I\!N}$ of piecewise linear convex functions such that for all $k$, $p^k \leq p$ and for some tolerance $\Delta_k > 0$,

$$p(x^{k+1}) - p^k(x^{k+1}) \leq \Delta_k.$$

Each approximate function $p^k$ will be obtained by using a bundle strategy in the sense that we build, step by step, piecewise linear convex functions

$\theta^1, \ldots, \theta^i, \ldots$ and we set $p^k = \theta^i$ when the solution $y^i$ of the unconstrained problem

$$(B_i^k) \quad \min\{\theta^i(x) + b(\nu_k, x) + \lambda_k^{-1}[\ K(x) - K(x^k) - \langle z^k, x - x^k\rangle\ ]\ \},$$

is such that $p(y^i) - \theta^i(y^i) \leq \Delta_k$. More precisely, the process is the following:

**Bundle scheme**

Let $x^k \in \text{int } C$ and let $\lambda_k, \Delta_k > 0$ be given. Compute $r(x^k) \in F(x^k)$.
Set $z^k = \nabla K(x^k) - \lambda_k r(x^k)$, $y^0 = x^k$ and $i = 1$.
**Step 1.** Choose a piecewise linear function $\theta^i \leq p$ and solve problem $(B_i^k)$ to obtain $y^i \in \text{int } C$.
**Step 2.** If $p(y^i) - \theta^i(y^i) \leq \Delta_k$, STOP and set $p^k = \theta^i$ and $x^{k+1} = y^i$. Otherwise increase $i$ by 1 and go to Step 1.

In order to prove that the STOP occurs after finitely many iterations, we have to impose conditions on the functions $\theta^i, i = 1, 2, \ldots$. Before presenting these conditions, first we observe that, by optimality of $y^i \in \text{int } C$, we have

$$\gamma^i \equiv \lambda_k^{-1}[z^k - \nabla K(y^i)] \in \partial[\theta^i + b(\nu_k, \cdot)](y^i). \tag{6.3}$$

Then we define the aggregate affine function $l^i$ by

$$l^i(y) = \theta^i(y^i) + b(\nu_k, y^i) + \langle \gamma^i, y - y^i\rangle, \quad y \in \text{int } C. \tag{6.4}$$

We have $l^i(y^i) = \theta^i(y^i) + b(\nu_k, y^i)$ and, using (6.3) and (6.4),

$$l^i(y) \leq \theta^i(y) + b(\nu_k, y), \quad \text{for all } y \in \text{int } C. \tag{6.5}$$

Now we require the following conditions on the functions $\theta^i$ for $i = 1, 2, \ldots$:

**(C1)** $\theta^i \leq p$,

**(C2)** $l^i \leq \theta^{i+1} + b(\nu_k, \cdot)$,

**(C3)** $p(y^i) + \langle s(y^i), \cdot - y^i\rangle \leq \theta^{i+1}$,

**(C4)** $p(y^0) + \langle s(y^0), \cdot - y^0\rangle \leq \theta^i$,

where $s(y^i)$ denotes a subgradient of $p$ at $y^i$. Here we suppose that, at each point of int $C$, one subgradient of $p$ is available.
The first three conditions are similar to those introduced in [35] in the framework of nonsmooth convex optimization. As in [35], they allow to prove that the STOP occurs in the bundle algorithm after finitely many iterations. Condition (C4) will be used in the next section to show the weak convergence of the sequence $\{x^k\}_{k\in I\!N}$ generated, step by step, by the bundle algorithm.

Let us now mention a few examples of functions $\theta^i$ satisfying Conditions (C1) to (C4). For the first function, we can take $\theta^1 = p(y^0) + \langle s(y^0), \cdot - y^0\rangle$ and for $i = 1, 2, \dots$, we can choose

$$\theta^{i+1} = \max_{0\le j\le i}\{p(y^j) + \langle s(y^j), \cdot - y^j\rangle\}. \tag{6.6}$$

It is easy to see that (C1), (C3) and (C4) are satisfied. Since $\theta^i \le \theta^{i+1}$, (C2) follows from (6.5). When $b(\nu_k, \cdot)$ is differentiable on int $C$, other choices are possible, for example,

$$\theta^{i+1} = \max_{0\le j\le i}\{\theta^i(y^i) + \langle \gamma^i - \nabla b(\nu_k, y^i), \cdot - y^i\rangle, p(y^j) + \langle s(y^j), \cdot - y^j\rangle\}. \tag{6.7}$$

Indeed, (C3) and (C4) are obvious and condition (C1) is satisfied because $\gamma^i - \nabla b(\nu_k, y^i) \in \partial\theta^i(y^i)$ and $s(y^i) \in \partial p(y^i)$. Finally, since

$$\theta^{i+1} \ge \theta^i(y^i) + \langle \gamma^i - \nabla b(\nu_k, y^i), \cdot - y^i\rangle = l^i - b(\nu_k, y^i) - \langle \nabla b(\nu_k, y^i), \cdot - y^i\rangle,$$

we have, using the subdifferential inequality, that

$$l^i \le \theta^{i+1} + b(\nu_k, y^i) + \langle \nabla b(\nu_k, y^i), \cdot - y^i\rangle \le \theta^{i+1} + b(\nu_k, \cdot),$$

i.e., condition (C2).

In the sequel we will also need to consider the following functions:

$$\begin{aligned}
\tilde{l}^i(y) &= l^i(y) + \lambda_k^{-1}\ [\ K(y) - K(x^k) - \langle z^k, y - x^k\rangle\ ],\\
\tilde{\theta}^i(y) &= \theta^i(y) + \lambda_k^{-1}\ [\ K(y) - K(x^k) - \langle z^k, y - x^k\rangle\ ].
\end{aligned}$$

Using (6.3) and (6.4), it is easy to see that, for all $y \in$ int $C$,

$$\tilde{l}^i(y) = \tilde{l}^i(y^i) + \lambda_k^{-1}\ [\ K(y) - K(y^i) - \langle \nabla K(y^i), y - y^i\rangle\ ]. \tag{6.8}$$

Moreover, we have

$$\tilde{\theta}^i(x^k) = \theta^i(x^k) \quad \text{and} \quad \tilde{l}^i(y^i) = \tilde{\theta}^i(y^i) + b(\nu_k, y^i), \tag{6.9}$$

and, by condition (C2),

$$\tilde{l}^i \leq \tilde{\theta}^{i+1} + b(\nu_k, \cdot). \tag{6.10}$$

**Proposition 6.1** *Suppose that $\partial p$ is bounded on bounded subsets of int $C$. If the stopping test is removed from the bundle algorithm and if the sequence $\{\theta^i\}_{i \in I\!N_0}$ satisfies conditions (C1) to (C3), then $p(y^i) - \theta^i(y^i) \to 0$.*

**Proof.** We proceed in three steps.

1. The sequence $\{\tilde{l}^i(y^i)\}_{i \in I\!N_0}$ is convergent and $y^{i+1} - y^i \to 0$.

   For all $i = 1, \ldots$ we have

$$\begin{aligned}
p(x^k) + b(\nu_k, x^k) &\geq \theta^{i+1}(x^k) + b(\nu_k, x^k) && \text{(by } (C1)) \\
&= \tilde{\theta}^{i+1}(x^k) + b(\nu_k, x^k) && \text{(by (6.9))} \\
&\geq \tilde{\theta}^{i+1}(y^{i+1}) + b(\nu_k, y^{i+1}) && \text{(definition of } y^{i+1}) \\
&= \tilde{l}^{i+1}(y^{i+1}) && \text{(by (6.9))} \\
&\geq \tilde{l}^i(y^{i+1}) && \text{(by (6.10))} \\
&= \tilde{l}^i(y^i) + \lambda_k^{-1}\, D_K(y^{i+1}, y^i) && \text{(by (6.8))},
\end{aligned}$$

   where $D_K(y, z) = K(y) - K(z) - \langle \nabla K(z), y - z \rangle \geq 0$ for all $y, z$. From these relations, we deduce that the sequence $\{\tilde{l}^i(y^i)\}_{i \in I\!N_0}$ is non-decreasing and bounded above by $p(x^k) + b(\nu_k, x^k)$. So it is convergent. Moreover, since $K$ is strongly convex of modulus $\beta > 0$, we also obtain that

$$\tilde{l}^{i+1}(y^{i+1}) - \tilde{l}^i(y^i) \geq \lambda_k^{-1}\, D_K(y^{i+1}, y^i) \geq (2\lambda_k)^{-1} \beta \|y^{i+1} - y^i\|^2 \geq 0.$$

   But then $y^{i+1} - y^i \to 0$ (strongly) because the left hand side tends to zero.

2. The sequence $\{y^i\}_{i \in I\!N}$ is bounded.

Let $y \in \text{int } C$ be fixed. Using successively (C1) and the definition of $\tilde{\theta}^{i+1}$, (6.10), (6.8) and the strong convexity of $K$, we have

$$
\begin{aligned}
p(y) + b(\nu_k, y) &+ \lambda_k^{-1} \, [ \, K(y) - K(x^k) - \langle z^k, y - x^k \rangle \, ] \\
&\geq \tilde{\theta}^{i+1}(y) + b(\nu_k, y) \\
&\geq \tilde{l}^i(y^i) + \lambda_k^{-1} D_K(y, y^i) \\
&\geq \tilde{l}^i(y^i) + (2\lambda_k)^{-1} \beta \|y - y^i\|^2.
\end{aligned}
$$

Since the sequence $\{\tilde{l}^i(y^i)\}_{i \in I\!N_0}$ is convergent, the sequence $\{y - y^i\}_{i \in I\!N}$ must be bounded and thus also the sequence $\{y^i\}_{i \in I\!N}$.

3. $p(y^{i+1}) - \theta^{i+1}(y^{i+1}) \to 0$.

   Using successively (C3), (C1) and the definition of the subgradient $s(y^{i+1})$, we obtain

$$
\begin{aligned}
\langle s(y^i), y^{i+1} - y^i \rangle &\leq \theta^{i+1}(y^{i+1}) - p(y^i) \\
&\leq p(y^{i+1}) - p(y^i) \\
&\leq \langle s(y^{i+1}), y^{i+1} - y^i \rangle.
\end{aligned}
$$

   Since the subdifferential $\partial p$ is bounded on the bounded sequence $\{y^i\}_{i \in I\!N}$, the sequence $\{s(y^i)\}_{i \in I\!N}$ is bounded and, as $\|y^{i+1} - y^i\| \to 0$, the opposite sides of the previous inequalities tend to zero. Hence

$$\theta^{i+1}(y^{i+1}) - p(y^i) \to 0 \quad \text{and} \quad p(y^{i+1}) - p(y^i) \to 0,$$

   such that

$$p(y^{i+1}) - \theta^{i+1}(y^{i+1}) = p(y^{i+1}) - p(y^i) + p(y^i) - \theta^{i+1}(y^{i+1}) \to 0.$$

This completes the proof. □

Since $\Delta_k > 0$, it follows from Proposition 6.1, that the STOP occurs after finitely many iterations in the bundle scheme. So $p^k$ is well defined and

$$p^k \leq p \quad \text{and} \quad p(x^{k+1}) - p^k(x^{k+1}) \leq \Delta_k. \tag{6.11}$$

Finally, the sequence $\{x^k\}_{k \in I\!N}$ generated by applying, step by step, the bundle algorithm is well defined and its convergence can be studied. It is the purpose of the next section.

**Remark 6.1** When $C \subseteq \text{int}(\text{dom } p)$, the subdifferential $\partial p(x)$ is nonempty on $C$ and there is no need to suppose that int $C$ is nonempty and to consider a barrier function $b(\nu_k, \cdot)$ in the subproblems $(B_i^k)$. In that case, $b(\nu_k, x)$ is replaced by $\Psi_C(x)$ in subproblem $(B_i^k)$. When $C$ is given by linear inequalities and $K$ is a strongly convex quadratic function as for example, $K = 1/2\|\cdot\|^2$, observe that subproblems $(B_i^k)$ become, in fact, convex quadratic programming problems.

## 6.2 Convergence of the Algorithm

We proceed in three steps to prove the convergence of the sequence $\{x_k\}_{k\in I\!N}$ generated by solving subproblems $(BSAP^k)$. First, we study the boundedness of $\{x_k\}_{k\in I\!N}$, then its weak convergence to some solution of problem $(GVIP)$ and finally its strong convergence. The results obtained are rather similar to those of Section 5.1 for $\varphi$ replaced by $p + \Psi_C$, dom $\varphi = C$ and $E = \text{int } C$.

Recall that the sequence $\{\lambda_k\}_{k\in I\!N}$ is chosen under the following form:

$$\begin{cases} \lambda_k = \mu_k/\eta_k, \forall k \in I\!N, \text{ with } \{\mu_k\}_{k\in I\!N} \text{ a sequence of positive numbers,} \\ \text{and } \eta_k = \begin{cases} \max\{1, \|r(x^0)\|\}, \text{ if } k = 0; \\ \max\{\eta_{k-1}, \|r(x^k)\|\}, \text{ if } k \geq 1. \end{cases} \end{cases}$$

Consider the sequence of Lyapounov functions $\{\Gamma^k(x^*, \cdot)\}_{k\in I\!N}$ given in (5.1) with $\varphi$ replaced by $p + \Psi_C$. The following lemma shows that we can obtain the same kind of upper bound on $\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k)$ as in Lemma 5.1.

**Lemma 6.1** *Assume that $F$ is a monotone multivalued mapping defined on $H$, that problem $(GVIP)$ admits at least one solution denoted by $x^*$, and that the following conditions are satisfied:*

**(i)** *$K : H \to I\!R$ is continuously differentiable and strongly convex with modulus $\beta > 0$ over dom $\varphi$;*

**(ii)** *$\nabla K$ is a Lipschitz continuous mapping with Lipschitz constant $\Lambda$ over dom $\varphi$;*

**(iii)** *$\{\mu_k\}_{k\in I\!N}$ is a nonincreasing sequence of positive numbers;*

**(iv)** *$p \in \Gamma_0(H)$ is such that $\partial p$ is bounded on bounded subsets of int $C$;*

**(v)** *$\{p^k\}_{k\in \mathbb{N}}, p \in \Gamma_0(H)$ are such that (6.11) is satisfied with $\Delta_k > 0$ such that $\sum_{k=0}^{+\infty} \Delta_k < +\infty$;*

**(vi)** *$\{b(\nu_k, \cdot)\}_{k\in \mathbb{N}}$ is a sequence of barrier functions associated with $C$, and there exists a sequence $\{w^k\}_{k\in \mathbb{N}}$ in int $C$ such that (6.2) holds.*

*Then, if $\{x^k\}_{k\in \mathbb{N}}$ denotes the sequence generated by solving subproblems $(BSAP^k)$, we have for all $k \in \mathbb{N}$,*

$$\begin{aligned}\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) &\le -c\|x^{k+1} - x^k\|^2 + \bar{T}^k + \mu_k^2 u \\ &\quad -(\mu_k/\eta_k)[\langle r(x^k), x^k - x^*\rangle + p(x^k) - p(x^*)],\end{aligned} \tag{6.12}$$

*with $c, u > 0$, $\bar{T}^k \ge 0$, and $\sum_{k=0}^{+\infty} \bar{T}^k < +\infty$.*

**Proof.** Let us review the proof of Lemma 5.1 with $\varphi$ replaced by $p + \Psi_C$, $\varphi^k$ replaced by $p^k + b(\nu_k, \cdot)$ and dom $\varphi$ by $C$.

The terms $s_1, s_{21}$ of inequality (5.7) can be treated in the same way. For the term $s_{22} + s_3$, the last line of (5.10)

$$\varphi^k(w^k) - \varphi(x^*) + \varphi(x^{k+1}) - \varphi^k(x^{k+1}),$$

is replaced by

$$\begin{aligned}&b(\nu_k, w^k) - \Psi_C(x^*) + \Psi_C(x^{k+1}) - b(\nu_k, x^{k+1}) \\ &+p^k(w^k) - p(x^*) + p(x^{k+1}) - p^k(x^{k+1}).\end{aligned}$$

For this expression, the assumptions $(v)$ and $(vi)$ imposed on $\{p^k\}_{k\in \mathbb{N}}$ and $\{b(\nu_k, \cdot)\}_{k\in \mathbb{N}}$ imply that

$$\Psi_C(x^{k+1}) - b(\nu_k, x^{k+1}) \le 0,$$

$$p(x^{k+1}) - p^k(x^{k+1}) \le \Delta_k, \text{ and}$$

$$p^k(w^k) - p(x^*) \le p(w^k) - p(x^*) \le \langle e^k, w^k - x^*\rangle \le \|e^k\| \, \|w^k - x^*\|,$$

where $e^k$ is any subgradient of $p$ at $w^k$ (it exists because $w^k \in$ int $C$). Moreover, since $\partial p$ is bounded on the bounded sequence $\{w^k\}_{k\in \mathbb{N}}$, there exists $d > 0$ such that for all $k$,

$$p^k(w^k) - p(x^*) \le d\,\|w^k - x^*\|.$$

So, in this case, the last line of (5.10) is lower or equal to

$$b(\nu_k, w^k) - \Psi_C(x^*) + \Delta_k + d\,\|w^k - x^*\|.$$

Gathering this with inequalities (5.7)–(5.13) and rearranging the terms, we obtain that inequality (6.12) holds with

$$\begin{array}{lcl} c & = & (1/2)(\beta - \tau - \gamma - \mu), \\ \bar{T}^k & = & \mu_0(1+d)\|w^k - x^*\| + \mu_0\, b(\nu_k, w^k) + (\Lambda^2/(2\tau))\|w^k - x^*\|^2 + \mu_0\Delta_k, \\ u & = & (1/(2\gamma)) + (1/(2\mu))\|r(x^*)\|^2, \\ \tau, & \gamma, & \mu > 0 \ \text{ such that } \tau + \gamma + \mu < \beta. \end{array}$$

Since the sequence $\{w^k\}_{k\in I\!N}$ has been chosen such that (6.2) holds and the series $\sum_{k=0}^{+\infty} \Delta_k$ is convergent, we have that $\sum_{k=0}^{+\infty} \bar{T}^k < +\infty$ and the proof is complete. □

When $H$ is a finite dimensional space, the assumption "$\partial p$ is bounded on bounded subsets of int $C$" is always true (see, for example, [106], Theorem 24.7).

We can then deduce the boundedness of the sequence $\{x^k\}_{k\in I\!N}$.

**Theorem 6.1** *Assume that all assumptions of Lemma 6.1 hold.*
*If* $\sum_{k=0}^{+\infty} \mu_k^2 < +\infty$, *then provided that* $x^0 \in C$, *the sequence* $\{x^k\}_{k\in I\!N}$ *is bounded. Moreover,*

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty, \ \textit{and}$$

$$\sum_{k=0}^{+\infty} (\mu_k/\eta_k)[\langle r(x^k), x^k - x^*\rangle + p(x^k) - p(x^*)] < +\infty.$$

**Proof.** Same proof as for Theorem 5.1. □

Now to ensure that each weak limit point of $\{x^k\}_{k\in I\!N}$ is a solution of problem $(GVIP)$, we use a gap function for problem $(GVIP)$ and Condition $(GZH_{x^*})$ just like in Section 5.1 with $\varphi$ replaced by $p + \Psi_C$, dom $\varphi$ replaced by $C$ and the set $E$ by int $C$. To be able to use successively Lemma 5.2 and Proposition 5.1, we have to prove the following proposition:

**Proposition 6.2** *(a) Assume that assumptions of Theorem 6.1 are satisfied as well as Condition* $(GZH_{x^*})$ *(i) and (ii). If* $F$ *is bounded on bounded subsets of int* $C$*, then* $\sum_{k=0}^{+\infty} \mu_k l(x^k) < +\infty$*.*
*(b) If* $\partial p$ *is bounded on bounded subsets of int* $C$ *and if there exists* $\delta_b > 0$ *such that*

$$b(\nu_k, x^k) - b(\nu_k, x^{k+1}) \leq \delta_b \|x^{k+1} - x^k\| \text{ for all } k, \tag{6.13}$$

*then there exists* $\delta > 0$ *such that, for all* $k$, $\|x^{k+1} - x^k\| \leq \delta\lambda_k$*.*
*(c) When* $C = \{x \mid g_i(x) \leq 0, i = 1, \ldots, m\}$ *with* $g_1, \ldots, g_m$ *convex functions from* $I\!R^n$ *to* $I\!R$*, inequality (6.13) is satisfied by the logarithmic barrier functions and by the inverse barrier functions provided that the barrier parameters be large enough.*

**Proof.** (a) See the proof of Proposition 5.5(a).

(b) From the optimality conditions applied to $x = x^k$, we obtain

$$\begin{aligned} \langle \nabla K(x^{k+1}) - \nabla K(x^k), x^{k+1} - x^k \rangle &\leq (\mu_k/\eta_k)[\langle r(x^k), x^k - x^{k+1} \rangle \\ &+ p^k(x^k) - p^k(x^{k+1}) + b(\nu_k, x^k) - b(\nu_k, x^{k+1})]. \end{aligned} \tag{6.14}$$

Since $K$ is strongly convex and $\|r(x^k)\| \leq \eta_k$, we derive from (6.14) that

$$\begin{aligned} \beta\|x^{k+1} - x^k\|^2 &\leq \mu_k \|x^{k+1} - x^k\| \\ &+ (\mu_k/\eta_k)[p^k(x^k) - p^k(x^{k+1}) + b(\nu_k, x^k) - b(\nu_k, x^{k+1})]. \end{aligned} \tag{6.15}$$

Now since $p^k \leq p$ and, by construction (see condition (C4)),

$$p^k(x) \geq p(x^k) + \langle s(x^k), x - x^k \rangle \text{ for all } x \in \text{int } C,$$

we have

$$\begin{aligned} p^k(x^k) - p^k(x^{k+1}) &\leq p(x^k) - p(x^k) - \langle s(x^k), x^{k+1} - x^k \rangle \\ &= \langle s(x^k), x^k - x^{k+1} \rangle \\ &\leq \|s(x^k)\| \, \|x^{k+1} - x^k\|. \end{aligned}$$

Since $\partial p$ is bounded on bounded subsets of int $C$, the sequence $\{\|s(x^k)\|\}_{k \in I\!N}$ is bounded and there exists $\delta_p > 0$ such that, for all $k$,

$$p^k(x^k) - p^k(x^{k+1}) \leq \delta_p \, \|x^{k+1} - x^k\|. \tag{6.16}$$

Finally, from (6.15), (6.16), (6.13), and since $\eta_k \geq 1$, we deduce that $\|x^{k+1} - x^k\| \leq \delta\mu_k$ for all $k$, with $\delta = (1/\beta)[\,1 + \delta_p + \delta_b\,]$.

(c) See Remark 5.1. □

We are now ready to state our main convergence result.

**Theorem 6.2** *Suppose that the following conditions are satisfied:*
*(a) Assumptions of Lemma 6.1 hold;*
*(b) Assumption $(GZH_{x^*})$ holds;*
*(c) $F$ is bounded on bounded subsets of int $C$;*
*(d) Inequality (6.13) holds;*
*(e) $\sum_{k=0}^{+\infty} \mu_k = +\infty,\ \sum \mu_k^2 < +\infty$.*
*Then the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded, $l(x^k) \to 0$ and any weak limit point of $\{x^k\}_k$ (and there exists at least one such point) is a solution of problem $(GVIP)$.*
*If, in addition, $\nabla K$ is weakly continuous on $C$, then $x^k \rightharpoonup \bar{x}$ and $\bar{x}$ is a solution of $(GVIP)$. If, in addition, the gap function $l$ is strongly convex on an open set containing $C$, then $x^k \to x^*$, the unique solution of $(GVIP)$. When $H$ is a finite dimensional space, the (strong) convergence of the whole sequence toward a solution is true under the only assumptions (a)–(e).*

**Proof.** The first part of the theorem follows immediately from Lemma 6.1, Theorem 6.1, Proposition 6.2, Lemma 5.2 and Proposition 5.1. The rest of the proof is the same as for Theorem 5.3. □

Using Propositions 5.2, 5.3 and 5.4 which give sufficient conditions to ensure that Assumption $(GZH_{x^*})$ is satisfied, we can particularize Theorem 6.2 to get more precise convergence theorems. However, before presenting them, and for the sake of simplicity, we collect in a statement, several assumptions used previously.

**Assumption $(A)$**:

*(i) Assumptions of Lemma 6.1 hold;*
*(ii) $\nabla K$ is weakly continuous on $C$;*
*(iii) Inequality (6.13) holds;*
*(iv) $\sum_{k=0}^{+\infty} \mu_k = +\infty,\ \sum_{k=0}^{+\infty} \mu_k^2 < +\infty$.*

Notice that when $H$ is a finite dimensional space, $(ii)$ is always true as also it is the case for assumption $(iv)$ of Lemma 6.1: $\partial p$ is bounded on bounded subsets of int $C$.

**Theorem 6.3** *Suppose that Assumption $(A)$ holds.*
*(a) If $F$ is paramonotone, weakly closed on $C$ and Lipschitz continuous on bounded subsets of int $C$, and if $F(x)$ is a bounded subset of $H$ for all $x \in C$, then the whole sequence $x^k \rightharpoonup \bar{x}$ where $\bar{x}$ is a solution of $(GVIP)$.*
*(b) If $F = \partial f$ with $f \in \Gamma_0(H)$ and $C \subseteq$ int(dom $f$), and if $\partial f$ is bounded on bounded subsets of int $C$, then the whole sequence $x^k \rightharpoonup \bar{x}$ where $\bar{x}$ is a solution of problem $(GVIP)$.*
*When $H$ is a finite dimensional space, the assumption on $\partial f$ is always true.*

**Proof.** By Theorem 6.2, it is sufficient to prove that Assumption $(GZH_{x^*})$ holds and that $F$ is bounded on bounded subsets of int $C$.
(a) Since $\partial p$ is bounded on bounded subsets of int $C$, it follows from Lemma 5.3 that $p$ is Lipschitz continuous on bounded subsets of int $C$. All the assumptions of Proposition 5.2 are then satisfied. Thus, Assumption $(GZH_{x^*})$ is satisfied. Finally, using Lemma 1.1, $F$ is bounded on bounded subsets of int $C$.
(b) By Lemma 5.3, $f$ and $p$ are Lipschitz continuous on bounded subsets of int $C$. So, using Proposition 5.3, Assumption $(GZH_{x^*})$ is satisfied. The conclusion follows because $F = \partial f$ is bounded on bounded subsets of int $C$.□

**Theorem 6.4** *Suppose that Assumption $(A)$ holds. If $F$ is strongly monotone on $C$ and bounded on bounded subsets of int $C$, then the whole sequence $x^k$ strongly converges to $x^*$, the unique solution of $(GVIP)$.*

**Proof.** From Proposition 5.4, we have that Assumption $(GZH_{x^*})$ is satisfied. Then the conclusion follows from Theorem 6.2 because $F$ is bounded on bounded subsets of int $C$ and the gap function $l(x) = \|x - x^*\|^2$ is strongly convex on $H$. □

**Remark 6.2** When $C \subseteq$ int (dom $p$), the subdifferential $\partial p(x)$ is nonempty on $C$ and there is no need to suppose that int $C$ is nonempty and to introduce a barrier function in subproblems $(BSAP^k)$. However, all our convergence results remain true in that case, provided that each assumption made on int $C$ be extended to $C$.

To conclude, note that we suppose that the sequence $\{p^k\}_{k\in\mathbb{N}}$ of functions approximating $p$, satisfies inequalities (6.11) where $\{\Delta_k\}_{k\in\mathbb{N}}$ is a sequence of positive numbers such that $\sum_{k=0}^{+\infty}\Delta_k < +\infty$. But practically, the choice of the sequence $\{\Delta_k\}_{k\in\mathbb{N}}$ remains an open issue. To be efficient, the tolerance $\Delta_k$ should be determined, not in advance, but once $x^k$ has been found by taking into account the behavior and the progress of the iterates to the solution. Such a strategy has been proposed in nonsmooth optimization, for solving nondifferentiable convex minimization problems (see [35]). In our context, this question deserves more investigations.

# Bibliography


[1] S. Adly, D. Goeleven, and M. Théra. Recession Mappings and Noncoercive Variational Inequalities. *Nonlinear Analysis, Theory, Methods and Applications*, 26:1573–1603, 1996.

[2] P. Alart and B. Lemaire. Penalization in Non–classical Convex Programming via Variational Convergence. *Mathematical Programming*, 51:307–331, 1991.

[3] Ya.I. Alber. Recurrence Relations and Variational Inequalities. *Soviet Mathematics, Doklady*, 27:511–517, 1983.

[4] Ya.I. Alber, A.N. Iusem, and M.V. Solodov. On the Projected Subgradient Method for Nonsmooth Convex Optimization in a Hilbert Space. *Mathematical Programming*, 81:23–35, 1998.

[5] H. Attouch. *Variational Convergence for Functions and Operators.* Pitman, London, 1984.

[6] H. Attouch, Z. Chbani, and A. Moudafi. Recession Operators and Solvability of Variational Problems in Reflexive Banach Spaces. *Series on Advances in Mathematics and Applied Sciences*, 18:51–67, 1994.

[7] H. Attouch and R.J.B. Wets. Approximation and Convergence in Nonlinear Optimization. In O.L. Mangasarian, R.R. Meyer, and S.M. Robinson, editors, *Nonlinear Programming 4*, pages 367–394. Academic Press, New York, New York, 1981.

[8] H. Attouch and R.J.B. Wets. Isometries for Legendre–Fenchel Transform. *Transactions of the American Mathematical Society*, 296:33–60, 1986.

[9] J.P. Aubin. *Applied Functional Analysis.* Wiley-Interscience, New York, 1979.

[10] J.P. Aubin and I. Ekeland. *Applied Nonlinear Analysis.* Wiley, New York, New York, 1984.

[11] G. Auchmuty. Variational Principles for Variational Inequalities. *Numerical Functional Analysis and Optimization*, 10:863–874, 1989.

[12] A. Auslender. *Optimisation. Méthodes Numériques.* Masson, Paris, 1976.

[13] A. Auslender, J.P. Crouzeix, and P. Fedit. Penalty–Proximal Methods in Convex Programming. *Journal of Optimization Theory and Applications*, 55:1–21, 1987.

[14] J.B. Baillon and G. Haddad. Quelques Propriétés des Opérateurs Angle–Bornés et n–Cycliquement Monotones. *Israel Journal of Mathematics*, 26:137–150, 1977.

[15] O. Barrientos. A Global Regularization Method for Solving the Finite Min–Max Problem. *Computational Optimization and Applications*, 11:277–295, 1998.

[16] D.P. Bertsekas and S.K. Mitter. A Descent Numerical Method for Optimization Problems with Nondifferentiable Cost Functionals. *SIAM Journal on Control*, 11:637–652, 1973.

[17] H. Brézis. Equations et Inéquations Non Linéaires dans les Espaces Vectoriels en Dualité. *Annales de l'Institut Fourier*, 18:115–175, 1968.

[18] H. Brézis. *Opérateurs Maximaux Monotones et Semi–Groupes de Contractions dans les Espaces de Hilbert.* North–Holland, Amsterdam, 1973.

[19] H. Brézis and A. Haraux. Images d'une Somme d'Opérateurs Monotones et Applications. *Israel Journal of Mathematics*, 23:165–186, 1976.

[20] A. Brønsted and R.T. Rockafellar. On the Subdifferentiability of Convex Functions. *Proceedings of the American Mathematical Society*, 16:605–611, 1965.

[21] M. Brohé. *Méthodes de Type Proximal Pour une Somme d' Opérateurs Maximaux Monotones.* PhD thesis, Université de Liège, 2000.

[22] F.E. Browder and W.V Petryshyn. Construction of Fixed Points of Nonlinear Mappings in Hilbert Space. *Journal on Mathematical Analysis and Applications*, 20:197–228, 1967.

[23] R.E. Jr. Bruck. An Iterative Solution of a Variational Inequality for Certain Monotone Operators in Hilbert Space. *Bulletin of the American Mathematical Society*, 81:890–892, 1975.

[24] R.S. Burachik and A.N. Iusem. A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Space. *SIAM Journal on Optimization*, 8:197–216, 1998.

[25] R.S. Burachik, A.N. Iusem, and B.F. Svaiter. Enlargement of Maximal Monotone Operators with Application to Variational Inequalities. *Set-Valued Analysis*, 5:159–180, 1997.

[26] R.S. Burachik, C.A. Sagastizábal, and B.F. Svaiter. Bundle Methods for Maximal Monotone Operators. In R. Tichatschke and M. Théra, editors, *Ill-Posed Variational Problems and Regularization Techniques, Lectures Notes in Economics and Mathematical Systems, No.477*, pages 49–64. Springer-Verlag, 1999.

[27] R.S. Burachik, C.A. Sagastizábal, and B.F. Svaiter. $\epsilon$-Enlargement of a Maximal Monotone Operator: Theory and Applications. In M. Fukushima and L. Qi, editors, *Reformulation-Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, pages 25–44. Kluwer Academic Publishers, 1999.

[28] Y. Censor, A.N. Iusem, and S. Zenios. An Interior Point Method with Bregman Functions for the Variational Inequality Problem with Paramonotone Operators. *Mathematical Programming*, 81:373–400, 1998.

[29] G.H.G. Chen and R.T. Rockafellar. Convergence Rates in Forward–Backward Splitting Methods. *SIAM Journal on Optimization*, 7:421–444, 1997.

[30] G. Cohen. Optimization by Decomposition and Coordination: A Unified Approach. *IEEE Transactions on Automatic Control*, AC–23:222–232, 1978.

[31] G. Cohen. Auxiliary Problem Principle and Decomposition of Optimization Problems. *Journal of Optimization Theory and Applications*, 32:277–305, 1980.

[32] G. Cohen. Nash Equilibria: Gradient and Decomposition Algorithms. *Large Scale Systems*, 12:173–184, 1987.

[33] G. Cohen. Auxiliary Problem Principle Extended to Variational Inequalities. *Journal of Optimization Theory and Applications*, 59:325–333, 1988.

[34] G. Cohen and D.L. Zhu. Decomposition Coordination Methods in Large Scale Optimization Problems: The Nondifferentiable case and the Use of Augmented Lagrangians. In J.B. Cruz, editor, *Advances in Large Scale Systems Theory and Applications 1*, pages 203–266. JAI Press, Greenwich, Connecticut, 1984.

[35] R. Correa and C. Lemaréchal. Convergence of Some Algorithms for Convex Minimization. *Mathematical Programming*, 62:261–275, 1993.

[36] J.P. Crouzeix. Pseudomonotone Variational Inequality Problems: Existence of Solutions. *Mathematical Programming*, 78:305–314, 1997.

[37] J.P. Crouzeix, P. Marcotte, and D. Zhu. Conditions Ensuring the Applicability of Cutting–plane Methods for Solving Variational Inequalities. *Mathematical Programming*, 88:521–539, 2000.

[38] J.P. Crouzeix, J.E. Martinez-Legaz, and M. Volle, editors. *Generalized Convexity, Generalized Monotonicity.* Nonconvex Optimization and its Applications, Vol. 27, Kluwer Academic Publishers, Dordrecht, 1998.

[39] J.C. Culioli and G. Cohen. Decomposition/Coordination Algorithms in Stochastic Optimization. *SIAM Journal on Control and Optimization*, 28:1372–1403, 1990.

[40] S. Dafermos. An Iterative Scheme for Variational Inequalities. *Mathematical Programming*, 26:40–47, 1983.

[41] A. Daniilidis and N. Hadjisavvas. On the Subdifferentials of Quasiconvex and Pseudoconvex Functions and Cyclic Monotonicity. *Journal on Mathematical Analysis and Applications*, 237:30–42, 1999.

[42] J. Eckstein. Approximate Iterations in Bregman–Function–Based Proximal Algorithms. *Mathematical Programming*, 83:113–124, 1998.

[43] J. Eckstein and D.P. Bertsekas. On the Douglas–Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators. *Mathematical Programming*, 55:293–318, 1992.

[44] I. Ekeland and R. Temam. *Convex Analysis and Variational Inequalities.* North–Holland, Amsterdam, 1976.

[45] R. El-Bachari. *Contribution à l'Etude des Algorithmes Proximaux: Décomposition et Perturbation Variationnelle.* PhD thesis, INSA de Rouen, 1996.

[46] N. El-Farouq. Pseudomonotone Variational Inequalities. Convergence of the Auxiliary Problem Method. Private Communication, 1997.

[47] N. El-Farouq and G. Cohen. Progressive Regularization of Variational Inequalities and Decomposition Algorithms. *Journal of Optimization Theory and Applications*, 97:407–433, 1998.

[48] S.C. Fang and E.L. Peterson. Generalized Variational Inequalities. *Journal of Optimization Theory and Applications*, 38:363–383, 1982.

[49] P. Fedit. *Contribution aux Méthodes Numériques en Programmation Mathématique Non Différentiable.* PhD thesis, Université de Clermont II, 1985.

[50] A.V. Fiacco and G. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, New York, New York, 1968.

[51] D. Gabay. Applications of the Method of Multipliers to Variational Inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, pages 299–331. North-Holland, Amsterdam, 1983.

[52] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and Operator–Splitting Methods in Nonlinear Mechanics.* SIAM Studies in Applied Mathematics, Philadelphia, Pennsylvania, 1989.

[53] D. Goeleven. On the Solvability of Noncoercive Linear Variational Inequalities in Separable Hilbert Space. *Journal of Optimization Theory and Applications*, 79:493–511, 1993.

[54] D. Goeleven and G. Salmon. Finite Element Analysis for Noncoercive Variational Inequalities. *Internal Report of the Department of Mathematics of the University of Namur*, 1995.

[55] D. Goeleven, G.E. Stavroulakis, G. Salmon, and P.D. Panagiotopoulos. Solvability Theory and Projection Methods for a Class of Singular Variational Inequalities: Elastostatic Unilateral Contact Applications. *Journal of Optimization Theory and Applications*, 95:263–293, 1997.

[56] P.T. Harker and J.S. Pang. Finite–Dimensional Variational Inequality and Nonlinear Complementarity Problems. A Survey of Theory, Algorithms and Applications. *Mathematical Programming*, 48:161–220, 1990.

[57] S. Haubruge. *Convergence Analysis and Applications of Some Splitting Methods for Finding a Zero of a Sum of Two Maximal Monotone Operators*. PhD thesis, University of Namur, 2000.

[58] J.L. Higle and S. Sen. Epigraphical Nesting: A Unifying Theory for the Convergence of Algorithms. *Journal of Optimization Theory and Applications*, 84:339–360, 1995.

[59] J.B. Hiriart-Urruty. $\epsilon$-Subdifferential Calculus. Convex Analysis and Optimization. *Notes in Mathematics*, 57:1–44, 1982.

[60] J.B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, Berlin, 1993.

[61] N. Huang. On the Generalized Implicit Quasivariational Inequalities. *Journal on Mathematical Analysis and Applications*, 216:197–210, 1997.

[62] G. Isac. *Complementarity Problems*. Lecture Notes in Mathematics No.1528, Springer-Verlag, Berlin, 1992.

[63] A.N. Iusem. On some Properties of Paramonotone Operators. *Journal of Convex Analysis*, 5:269–278, 1998.

[64] A.A. Kaplan. On Convex Programming with Internal Regularization. *Soviet Mathematics, Doklady Akademii Nauk*, 19:795–799, 1978.

[65] A.A. Kaplan. Convex Programming Algorithms Using Smoothing of Exact Penalty Functions. *Siberian Mathematical Journal*, 23:491–500, 1982.

[66] A.A. Kaplan and R. Tichatschke. Auxiliary Problem Principle and the Approximation of Variational Inequalities with Non–Symmetric Multi–Valued Operators. In M. Théra, editor, *Canadian Mathematical Society Conference Proceedings Series, No.27*, pages 185–209, 2000.

[67] S. Karamardian. Generalized Complementarity Problem. *Journal of Optimization Theory and Applications*, 8:161–168, 1971.

[68] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Application*. Academic Press, New York, New York, 1980.

[69] K.C. Kiwiel. Proximal Minimization Methods with Generalized Bregman Functions. *SIAM Journal on Control and Optimization*, 35:1142–1168, 1997.

[70] I. Konnov. On Quasimonotone Variational Inequalities. *Journal of Optimization Theory and Applications*, 99:165–181, 1998.

[71] I. Konnov. *Combined Relaxation Methods for Variational Inequalities*. Springer-Verlag, Berlin, 2001.

[72] I. Konnov and J.C. Yao. On the Generalized Vector Variational Inequality Problem. *Journal on Mathematical Analysis and Applications*, 206:42–58, 1997.

[73] B. Lemaire. Coupling Optimization Methods and Variational Convergence. In K.H. Hoffmann, J.B. Hiriart-Urruty, C. Lemaréchal, and J. Zowe, editors, *Trends in Mathematical Optimization, International Series of Numerical Mathematics*, pages 163–179. Birkhauser Verlag, Basel, 1988.

[74] B. Lemaire. On the Convergence of Some Iterative Methods for Convex Minimization. In R. Durier and C. Michelot, editors, *Recent Developments in Optimization, Lecture Notes in Economics and Mathematical Systems*, volume 429, pages 252–268. Springer-Verlag, Berlin, 1995.

[75] C. Lemaréchal, A. Nemirovskii, and Yu. Nesterov. New Variants of Bundle Methods. *Mathematical Programming*, 69:111–148, 1995.

[76] C. Lemaréchal, J.J. Strodiot, and A. Bihain. On a Bundle Method for Nonsmooth Optimization. In O.L. Mangasarian, R.R. Meyer, and S.M. Robinson, editors, *Nonlinear Programming 4*, pages 245–282. Academic Press, 1981.

[77] P.L. Lions and B. Mercier. Splitting Algorithms for the Sum of Two Nonlinear Operators. *SIAM Journal on Numerical Analysis*, 16:964–979, 1979.

[78] Z. Luo and P. Tseng. A Decomposition Property for a Class of Square Matrices. *Applied Mathematical Letters*, 4:67–69, 1991.

[79] T.L. Magnanti and G. Perakis. Averaging Schemes for Variational Inequalities and Systems of Equations. *Mathematics of Operations Research*, 22:568–587, 1997.

[80] T.L. Magnanti and G. Perakis. The Orthogonality Theorem and the Strong–f–Monotonicity Condition for Variational Inequality Algorithms. *SIAM Journal on Optimization*, 7:248–273, 1997.

[81] P. Mahey, V.H. Nguyen, and D.T. Pham. Proximal Techniques for the Decomposition of Convex Programs. *Rapport de Recherches Laboratoire ARTEMIS*, RR 87-M-:21–37, 1992.

[82] S. Makler-Scheimberg, V.H. Nguyen, and J-J. Strodiot. Family of Perturbation Methods for Variational Inequalities. *Journal of Optimization Theory and Applications*, 89:423–452, 1996.

[83] P. Marcotte and J.H. Wu. On the Convergence of Projection Methods: Application to the Decomposition of Affine Variational Inequalities. *Journal of Optimization Theory and Applications*, 85:347–362, 1995.

[84] B. Martinet. Régularisation d' Inéquations Variationnelles par Approximations Successives. *Revue Française d'Informatique et de Recherche Opérationelle*, 4(R–3):154–158, 1970.

[85] M.A. Mataoui. *Contributions à la Décomposition et à l'Agrégation des Problèmes Variationnels.* PhD thesis, Ecole des Mines de Paris, Fontainebleau, 1990.

[86] J. Mawhin and M. Willem. *Critical Point Theory and Hamiltonian Systems.* Springer-Verlag, New York, New York, 1989.

[87] U. Mosco. Convergence of Convex Sets and of Solutions of Variational Inequalities. *Advances in Mathematics*, 3:510–585, 1969.

[88] K. Mouallif. Sur la Convergence d'une Méthode Associant Pénalisation et Régularisation. *Bulletin de la Société Royale des Sciences de Liège*, 56:175–180, 1987.

[89] K. Mouallif, V.H. Nguyen, and J.J. Strodiot. A Perturbed Parallel Decomposition Method for a Class of Nonsmooth Convex Minimization Problems. *SIAM Journal on Control and Optimization*, 29:829–847, 1991.

[90] K. Mouallif and P. Tossings. Une Méthode de Pénalisation Exponentielle Associée à une Régularisation Proximale. *Bulletin de la Société Royale des Sciences de Liège*, 56:181–190, 1987.

[91] A. Moudafi. Coupling Proximal Methods and Variational Convergence. *Zeitschrift fur Operations Research*, 38:269–280, 1993.

[92] A. Moudafi and M.A. Noor. New Convergence Results of Iterative Methods for Set-Valued Mixed Variational Inequalities. *Mathematical Inequalities and Applications*, 3:295–303, 2000.

[93] A. Moudafi and M. Théra. Finding a Zero of the Sum of Two Maximal Monotone Operators. *Journal of Optimization Theory and Applications*, 94:425–448, 1997.

[94] Yu.E. Nesterov and A.S. Nemirovskii. *Interior–Point Polynomial Algorithms in Convex Programming.* SIAM Publications, Philadelphia, Pennsylvania, 1994.

[95] E.A. Nurminski. $\epsilon$-Subgradient Mapping and the Problem of Convex Optimization. *Cybernetics*, 21:796–800, 1986.

[96] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York, New York, 1970.

[97] P.D. Panagiotopoulos. *Inequality Problems in Mechanics and Applications: Convex and Nonconvex Energy Functions.* Birkhauser Verlag, Basel, 1985.

[98] P.D. Panagiotopoulos and G.E. Stavroulakis. New Types of Variational Principles Based on the Notion of Quasidifferentiability. *Acta Mechanics*, 94:171–194, 1994.

[99] J.S. Pang and D. Chan. Iterative Methods for Variational and Complementarity Problems. *Mathematical Programming*, 24:284–313, 1982.

[100] D. Pascali and S. Sburlan. *Nonlinear Mappings of Monotone Type.* Sijthoff and Noordhoff International Publishers, Alphen aan den Rijn, The Netherlands, 1978.

[101] G.B. Passty. Ergodic Convergence to a Zero of the Sum of Monotone Operators in Hilbert Space. *Journal on Mathematical Analysis and Applications*, 72:383–390, 1979.

[102] M. Patriksson. *Nonlinear Programming and Variational Inequality Problems.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[103] B.T. Polyak. *Introduction to Optimization.* Optimization Software, New York, New York, 1987.

[104] M. Qian. *Variable Metric Proximal Point Algorithm: Convergence Theory and Applications.* PhD thesis, University of Washington, 1992.

[105] A. Renaud and G. Cohen. An Extension of the Auxiliary Problem Principle to Nonsymmetric Auxiliary Operators. *ESAIM: Control, Optimisation and Calculus of Variations*, 2:281–306, 1997.

[106] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[107] R.T. Rockafellar. On the Maximal Monotonicity of Subdifferential Mapping. *Pacific Journal of Mathematics*, 33:209–216, 1970.

[108] R.T. Rockafellar. On the Maximality of Sums of Nonlinear Monotone Operators. *Transactions of the American Mathematical Society*, 149:75–88, 1970.

[109] R.T. Rockafellar. Monotone Operators and the Proximal Point Algorithm. *SIAM Journal on Control and Optimization*, 14:877–898, 1976.

[110] R. Saigal. Extension of the Generalized Complementarity Problem. *Mathematics of Operations Research*, 1:260–266, 1976.

[111] G. Salmon, V.H. Nguyen, and J-J. Strodiot. A Perturbed and Inexact Version of the Auxiliary Problem Method for solving General Variational Inequalities with a Multivalued Operator. In V.H. Nguyen, J-J. Strodiot, and P. Tossings, editors, *Optimization, Lecture Notes in Economics and Mathematical Systems, 481*, pages 396–418. Springer–Verlag, Berlin, 2000.

[112] G. Salmon, V.H. Nguyen, and J-J. Strodiot. Coupling the Auxiliary Problem Principle and the Epiconvergence Theory to Solve General Variational Inequalities. *Journal of Optimization Theory and Applications*, 104:629–657, 2000.

[113] G. Salmon, V.H. Nguyen, and J-J. Strodiot. A Pertubed Auxiliary Problem Method for Paramonotone Multivalued Mappings. In N. Hadjisavvas and P. Pardalos, editors, *Advances in Convex Analysis and Global Optimization, Non Convex Optimization and Applications Series.* Kluwer Academic Publishers, 2001. To Appear.

[114] G. Salmon, J-J. Strodiot, and V.H. Nguyen. A Bundle Method to Solve Multivalued Variational Inequalites. Submitted to SIAM Journal on Optimization, 2001.

[115] H. Schramm and J. Zowe. A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results. *SIAM Journal on Optimization*, 2:121–152, 1992.

[116] F. Sharifi-Mokhtarian and J.L. Goffin. Long-Step Interior-Point Algorithms for a Class of Variational Inequalities with Monotone Operators. *Journal of Optimization Theory and Applications*, 97:181–210, 1998.

[117] M.H. Shih and K.K. Tan. Browder-Hartman-Stampacchia Variational Inequalities for Multi-Valued Monotone Operators. *Journal on Mathematical Analysis and Applications*, 134:431–440, 1988.

[118] M.V. Solodov and B.F. Svaiter. A Hybrid Approximate Extragradient-Proximal Point Algorithm Using the Enlargement of a Maximal Monotone Operator. *Set-Valued Analysis*, 7:323–345, 1999.

[119] M.V. Solodov and S.K. Zavriev. Error-Stability Properties of Generalized Gradient-type Algorithms. *Journal of Optimization Theory and Applications*, 98:663–680, 1998.

[120] Y. Sonntag. *Convergence au Sens de Mosco: Théorie et Applications à l'Approximation des Solutions d'Inéquations*. PhD thesis, Université de Provence, 1982.

[121] J.J. Strodiot and V.H. Nguyen. On the Numerical Treatment of the Inclusion $0 \in \partial f(x)$. In J.J. Moreau, P.D. Panagiotopoulos, and G. Strang, editors, *Topics in Nonsmooth Mechanics*, pages 267–294. Birkhauser Verlag, 1988.

[122] J.J. Strodiot, V.H. Nguyen, and N. Heukemes. $\epsilon-$Optimal Solutions in Nondifferentiable Convex Programming and Some Related Questions. *Mathematical Programming*, 25:307–328, 1983.

[123] K. Taji, M. Fukushima, and T. Ibaraki. A Globally Convergent Newton Method for Solving Strongly Monotone Variational Inequalities. *Mathematical Programming*, 58:369–383, 1993.

[124] M. Teboulle. Convergence of Proximal–Like Algorithms. *SIAM Journal on Optimization*, 7:1069–1083, 1997.

[125] P. Tossings. *Sur les Zéros des Opérateurs Maximaux Monotones et Applications*. PhD thesis, Université de Liège, 1990.

[126] P. Tseng. Further Applications of a Splitting Algorithm to Decomposition in Variational Inequalities and Convex Programming. *Mathematical Programming*, 48:249–263, 1990.

[127] P. Tseng. Applications of a Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities. *SIAM Journal on Control and Optimization*, 29:119–138, 1991.

[128] M. Willem. *Analyse Convexe et Optimization*. Ciaco, Bruxelles, 1989.

[129] S.J. Wright. *Primal–Dual Interior Point Methods*. SIAM, Philadelphia, Pennsylvania, 1997.

[130] J.H. Wu. Modified Primal Path-Following Scheme for the Monotone Variational Inequality Problem. *Journal of Optimization Theory and Applications*, 95:189–208, 1997.

[131] J.H. Wu. Long-Step Primal Path-Following Algorithm for Monotone Variational Inequality Problems. *Journal of Optimization Theory and Applications*, 99:509–531, 1998.

[132] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*. Springer-Verlag, New York, New York, 1990.

[133] C. Zhu. Asymptotic Convergence Analysis of the Forward-Backward Splitting Algorithm. *Mathematics of Operations Research*, 20:449–464, 1995.

[134] D. Zhu. The Decomposition Method for the Variational Inequalities with Multivalued Mapping. Private Communication, 1999.

[135] D. Zhu and P. Marcotte. New Classes of Generalized Monotonicity. *Journal of Optimization Theory and Applications*, 87:457–471, 1995.

[136] D. Zhu and P. Marcotte. Co–Coercivity and its Role in the Convergence of Iterative Schemes for Solving Variational Inequalities. *SIAM Journal on Optimization*, 6:714–726, 1996.